

A nonparametric probabilistic approach for quantifying uncertainties in low-dimensional and high-dimensional nonlinear models

Christian Soize, Charbel Farhat

▶ **To cite this version:**

Christian Soize, Charbel Farhat. A nonparametric probabilistic approach for quantifying uncertainties in low-dimensional and high-dimensional nonlinear models. International Journal for Numerical Methods in Engineering, 2017, 109, pp.837-888. 10.1002/nme.5312 . hal-01353194

HAL Id: hal-01353194
https://hal.science/hal-01353194

Submitted on 10 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Preprint, International Journal for Numerical Methods in Engineering, May 2016

A nonparametric probabilistic approach for quantifying uncertainties in low- and high-dimensional nonlinear models

C. Soize[1]*and C. Farhat[2]

[1] *Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, Université Paris-Est, 5 bd Descartes, 77454 Marne-la-Vallee, France*
[2] *Department of Aeronautics and Astronautics, Department of Mechanical Engineering, and Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA*

SUMMARY

A nonparametric probabilistic approach for modeling uncertainties in projection-based, nonlinear, reduced-order models is presented. When experimental data is available, this approach can also quantify uncertainties in the associated high-dimensional models. The main underlying idea is two-fold. First, to substitute the deterministic Reduced-Order Basis (ROB) with a stochastic counterpart. Second, to construct the probability measure of the Stochastic Reduced-Order Basis (SROB) on a subset of a compact Stiefel manifold in order to preserve some important properties of a ROB. The stochastic modeling is performed so that the probability distribution of the constructed SROB depends on a small number of hyperparameters. These are determined by solving a reduced-order statistical inverse problem. The mathematical properties of this novel approach for quantifying model uncertainties are analyzed through theoretical developments and numerical simulations. Its potential is demonstrated through several example problems from computational structural dynamics. Accepted for publication in International Journal for Numerical Methods in Engineering, 30 May 2016

Received ...

KEY WORDS: modeling errors; model uncertainties; nonparametric stochastic approach; reduced-order model; model order reduction; uncertainty quantification

Notation

Throughout this paper:

A real, deterministic variable is denoted by a lower case letter such as y.
A real, deterministic vector is denoted by a boldface, lower case letter such as in $\mathbf{y} = (y_1, \ldots, y_N)$.
A real, random variable is denoted by an upper case letter such as Y.
A real, random vector is denoted by a boldface, upper case letter such as in $\mathbf{Y} = (Y_1, \ldots, Y_N)$.
A real, deterministic matrix is denoted by an upper (or lower) case letter between brackets such as $[A]$ (or $[a]$).
A real, random matrix is denoted by a boldface, upper case letter between brackets such as $[\mathbf{A}]$.

$\|\mathbf{y}\|$ designates the Euclidean norm of vector $\mathbf{y}$.
$< \mathbf{y}, \mathbf{z} >$ designates the Euclidean inner product of $\mathbf{y}$ and $\mathbf{z}$.
E designates the mathematical expectation.
$\mathbb{M}_{N,n}$ denotes the set of $N \times n$ real matrices.
$\mathbb{M}_n$ denotes the set of square $n \times n$ real matrices.

*Correspondence to: Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, Université Paris-Est, 5 bd Descartes, 77454 Marne-la-Vallee, France. E-mail: christian.soize@univ-paris-est.fr

$\mathbb{M}_n^{\mathrm{S}}$ denotes the set of symmetric $n \times n$ real matrices.
$\mathbb{M}_n^{\mathrm{SS}}$ denotes the set of skew-symmetric $n \times n$ real matrices.
$\mathbb{M}_N^{+}$ denotes the set of Symmetric Positive-Definite (SPD) $N \times N$ real matrices.
$\mathbb{M}_N^{+0}$ denotes the set of symmetric positive $N \times N$ real matrices.
$\mathbb{M}_N^{\mathrm{u}}$ denotes the set of upper triangular $N \times N$ real matrices with strictly positive diagonal entries.
A_{jk} designates the entry $[A]_{jk}$ of matrix $[A]$.
tr$\{[A]\}$ designates the trace of matrix $[A]$.
$[A]^T$ designates the transpose of matrix $[A]$.
$\|A\|_F$ designates the Frobenius norm of matrix $[A]$, with $\|A\|_F^2 = \mathrm{tr}\{[A]^T\,[A]\}$.
$\|A\|_M$ is defined by $\|A\|_M^2 = \mathrm{tr}\{[A]^T[M]\,[A]\}$, where $[M]$ is a positive definite matrix.
$[I_n]$ denotes the identity matrix in $\mathbb{M}_n$.
$[0_{N,n}]$ denotes the zero matrix in $\mathbb{M}_{N,n}$.
δ_{jk} denotes Kronecker's symbol and therefore verifies $\delta_{jk} = 0$ if $j \neq k$ and $\delta_{jk} = 1$ if $j = k$.
$\mathbb{1}_B(\mathbf{x})$ designates the indicator function of a set B defined by $\mathbb{1}_B(\mathbf{x}) = 1$ if $\mathbf{x} \in B$ and $\mathbb{1}_B(\mathbf{x}) = 0$ if $\mathbf{x} \notin B$.
i denotes the pure imaginary complex number satisfying $\mathrm{i}^2 = -1$.

1. INTRODUCTION

1.1. Background: high-dimensional and projection-based reduced-order models

Today, the potential of physics-based (or Partial Differential Equation (PDE)-based), High-Dimensional computational Models (HDMs) for providing deeper understanding of complex phenomena, enhancing system performance, and predicting the unknown is recognized in almost every field of science and engineering [1]. However, in many computational mechanics applications, nonlinear, time-dependent numerical simulations based on HDMs remain so computationally intensive or cost-prohibitive that they cannot be used as often as needed, or are more often used in special circumstances than routinely. For this reason, nonlinear, projection-based Model Order Reduction (MOR) has recently emerged as a promising if not indispensable numerical tool for parametric applications such as, to name only a few, design, design optimization, statistical analysis, and simulation-based decision making [2, 3, 4, 5, 6, 7, 8, 9, 10, 11].

In general, a $\boldsymbol{\mu}$-parametric, high-fidelity, physics-based computational model is high-dimensional because its underlying spatial discretization is performed *a priori* — that is, before any significant knowledge about the response of the system to be analyzed is developed. On the other hand, the projection of such a computational model of dimension N onto a subspace of low dimension $n << N$ represented by an associated $(N \times n)$ Reduced-Order Basis (ROB) $[V]$ leads to a Reduced-Order Model (ROM) of much lower dimension n. When $[V]$ is carefully constructed *a posteriori* — that is, after some knowledge about the response of the system has been developed — the corresponding ROM can capture the dominant behavior of the underlying $\boldsymbol{\mu}$-parametric HDM and therefore retain most of its fidelity. In general, knowledge about the system response is obtained during a *training* procedure that is performed *offline*. During this procedure, the model parameters represented here by the parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{N_p})$ belonging to the parameter space $\mathcal{C}_{\boldsymbol{\mu}}$ are sampled at a *few* points using a greedy but effective sampling strategy (for example, see [12]), and a set of problems related to the main problem of interest are solved to obtain a set of parametric solution snapshots. Then, these snapshots are compressed using, for example, the Singular Value Decomposition (SVD) to construct a *global* ROB. In general, the sampling strategy is designed so that the global ROB is reliable in a large region of the model parameter domain. Unfortunately, despite its low dimension, the resulting global (or $\boldsymbol{\mu}$-parametric) ROM does not necessarily guarantee computational feasibility. This is because the construction of this projection-based ROM does not scale only with its size n, but also with that of the underlying HDM, $N >> n$. In the deterministic setting, this issue is particularly problematic for nonlinear problems because the ROM needs to be repeatedly reconstructed to address, for example, time-dependency or Newton iterations for implicit solution strategies. This caveat is remedied by equipping an MOR method

with a rigorous procedure for approximating the resulting reduced operators whose computational complexity scales only with the small size n of the ROM [2, 4, 13, 14]. Such a procedure is also known in the literature as hyper reduction [15]. It transforms the nonlinear ROM into a hyper reduced ROM that guarantees feasibility, while maintaining as much as possible a desired level of accuracy.

From the above discussion, it follows that a nonlinear ROM or hyper reduced ROM inherits the modeling errors and associated uncertainties of its underlying HDM — including model form uncertainties. It is also tainted by additional errors introduced by the reduction processes highlighted above. Hence, if MOR is essential for enabling simulation-based decision making, Uncertainty Quantification (UQ) for ROMs is critical for certifying the decisions they enable.

1.2. Output-predictive error method

For a given HDM, several methods are currently available for analyzing model uncertainties induced by modeling errors, including model form uncertainties. The most popular one is the standard output-predictive error method introduced in [16]. In the context of ROMs however, this method has a major drawback. For example, it is not well-suited for design optimization problems where ROMs are particularly needed, because it does not enable the $\boldsymbol{\mu}$-parametric HDM and its associated ROM to learn from data.

1.3. Parametric probabilistic methods for modeling uncertainties

An alternative family of methods for analyzing model uncertainties is the familty of parametric probabilistic methods for UQ. This approach is relatively well developed for modeling model parameter uncertainty, at least for a reasonably small number of parameters. It consists in constructing prior and posterior stochastic models of uncertain model parameters pertaining, for example, to geometry, boundary conditions, material properties $\cdots$ [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]. This approach was shown to be computationally efficient for both a $\boldsymbol{\mu}$-parametric HDM and its associated $\boldsymbol{\mu}$-parametric ROM (for example, see [28, 5, 29]), and for large-scale statistical inverse problems [30, 31, 32, 33, 34, 35, 36, 37]. However, it does not take into account neither the model uncertainties induced by modeling errors introduced during the construction of a $\boldsymbol{\mu}$-parametric HDM (model form uncertainties), nor those due to model reduction [27, 29, 38].

1.4. Nonparametric probabilistic approach for modeling uncertainties

A nonparametric probabilistic approach for modeling uncertainties due to more general modeling errors was introduced in [38, 39, 40, 41], in the context of *linear* structural dynamics. It is organized in two steps. In the first one, a linear ROM of dimension n is constructed using a linear HDM with N degrees of freedom (dofs) — that is, of dimension N — and a linear, projection-based MOR. In the second step, a linear Stochastic ROM (SROM) is constructed by substituting the deterministic matrices underlying the linear ROM with random matrices for which the probability distributions are constructed using the Maximum Entropy (MaxEnt) principle [42, 43, 44]. Specifically, the construction of the linear SROM is performed under constraints generated from available information such as algebraic properties (positiveness, integrability of the inverse, $\cdots$) and statistical information (for example, the equality between mean and nominal values). This nonparametric probabilistic approach was extended for different ensembles of random matrices [40, 45] and linear boundary value problems [46]. It was also experimentally validated and applied for linear problems in composites [47], viscoelasticity [48, 49], dynamic substructuring [50, 51, 52, 53, 54], vibroacoustics [48, 55, 56, 57], soil-structure interaction and earthquake engineering [58, 59, 60], and robust design and optimization [61, 62]. More recently, it was further extended to account for some nonlinear geometrical effects in structural analysis [63, 64]. However, the latter extension is strongly related to the mathematical properties of the nonlinear elasticity operator considered in [63, 64]. *A priori*, it does not hold for arbitrarily nonlinear systems.

1.5. Main objective and organization of this paper

For all reasons outlined above, given a $\boldsymbol{\mu}$-parametric nonlinear ROM and corresponding experimental data, the main objective of this paper is to present a nonparametric probabilistic approach for modeling the uncertainties due to the following two types of modeling errors responsible for discrepancies between predictions performed using this nonlinear ROM and corresponding experimental data:

- The modeling errors associated with *nonlinear* MOR in general, independently of the type or source of nonlinearities. In this regard, it is noted here that when the dimension n of the $\boldsymbol{\mu}$-parametric nonlinear ROM approaches N, and the number m_μ of parameter vectors $\boldsymbol{\mu}$ that are sampled in $\mathcal{C}_{\boldsymbol{\mu}}$ for constructing the global ROB $[V]$ approaches infinity, the modeling errors of this type approach zero.
- Those introduced in the construction of the underlying $\boldsymbol{\mu}$-parametric nonlinear HDM, including model form errors.

From the second bullet above, it follows that the nonparametric probabilistic approach presented in this paper can also quantify uncertainties in any HDM for which a ROM can be constructed.

Furthermore, it is well-known that many applications such as robust design optimization call for a $\boldsymbol{\mu}$-parametric computational model that is capable of reproducing the typical experimental variability induced by a manufacturing process, and/or an imperfect if not noisy set of measurements. Therefore, a related objective of this paper is to ensure that the proposed nonparametric probabilistic approach can also account for experimental variability.

To this effect, the remainder of this paper is organized as follows. In Section 2, the proposed nonparametric probabilistic approach for modeling uncertainties is described using a simple example that introduces and highlights its key components. In Section 3, a method for constructing the associated Stochastic ROB (SROB) is presented. Its underlying theory is given in a discrete form that is readily applicable to computational models. A continuous interpretation of this theory is provided in Appendix D. In Section 4, the potential of the SROB for taking into account modeling errors is analyzed through the discussion of a simple numerical example. In particular, the performance of the proposed approach for identifying the hyperparameters of the probability model underlying the constructed SROB is assessed. Section 5 is devoted to a computational structural dynamics application of the proposed approach to UQ in nonlinear MOR that is simple to reproduce by the interested reader. Section 6 presents yet another example problem which focuses on the assessment of the ability of the $\boldsymbol{\mu}$-parametric SROM to account for the errors induced by modeling a system using a $\boldsymbol{\mu}$-parametric ROM instead of the underlying $\boldsymbol{\mu}$-parametric HDM. Finally, Section 7 concludes this paper.

2. NONPARAMETRIC PROBABILISTIC APPROACH FOR UNCERTAINTY ANALYSIS OF NONLINEAR MODEL ORDER REDUCTION METHODS

To begin, the proposed approach for modeling the uncertainties due to both model form and model reduction errors and its main underlying idea are outlined here, before their theoretical and algorithmic underpinnings are presented in the remainder of this paper. For this purpose, the context is set to that of a nonlinear Finite Element (FE) structural dynamics model. However, the overall approach is equally applicable, at least in principle, to any $\boldsymbol{\mu}$-parametric, nonlinear, computational model. Essentially, given such a model and a Quantity of Interest (QoI), a $\boldsymbol{\mu}$-parametric, nonlinear, projection-based ROM is constructed using a standard procedure such as that outlined in Section 1. Next, an SROB that verifies a specific set of mathematical properties is generated by randomizing the deterministic ROB underlying the nonlinear ROM. This SROB depends on a vector-valued hyperparameter that is identified by formulating an appropriate cost function and solving an associated optimization problem. Finally, the $\boldsymbol{\mu}$-parametric SROM associated with the global SROB is computed using the same projection approach chosen for constructing the deterministic nonlinear ROM.

2.1. Parametric nonlinear high-dimensional computational model

Consider the $\boldsymbol{\mu}$-parametric, nonlinear, computational model

$$[M]\,\ddot{\mathbf{y}}(t) + \mathbf{g}(\mathbf{y}(t), \dot{\mathbf{y}}(t); \boldsymbol{\mu}) = \mathbf{f}(t; \boldsymbol{\mu}) \quad , \quad t \in]t_0, T] \tag{2.1}$$

defined on $\mathbb{R}^N$, arising from the *large-scale* FE semi-discretization of the PDEs governing the dynamic equilibrium of a given structure, equipped with the initial conditions

$$\mathbf{y}(t_0) = \mathbf{y}_0 \quad , \quad \dot{\mathbf{y}}(t_0) = \mathbf{y}_1 \,,$$

and subject to $N_{\text{CD}} < N$ linear constraints of interest written in matrix form as follows

$$[B]^T \mathbf{y}(t) = \mathbf{0}_{N_{\text{CD}}} \quad , \quad t \in [t_0, T] \,, \tag{2.2}$$

where t_0 and T are two given time-instances satisfying $-\infty < t_0 < T < +\infty$, and $\mathbf{y}_0$ and $\mathbf{y}_1$ are two given vectors in $\mathbb{R}^N$ satisfying the specified constraints.

In Eq. (2.1) above, $\boldsymbol{\mu}$ is as before the vector parameter belonging to $\mathcal{C}_{\boldsymbol{\mu}} \subset \mathbb{R}^{m_\mu}$, t denotes time, $\mathbf{y}(t) = (y_1(t), \ldots, y_N(t))$ is the $\mathbb{R}^N$ vector of the N displacement dofs, $\dot{\mathbf{y}}(t) = d\mathbf{y}(t)/dt$ and $\ddot{\mathbf{y}}(t) = d^2\mathbf{y}(t)/dt^2$ are the corresponding velocity and acceleration vectors, $[M]$ is the mass matrix belonging to $\mathbb{M}_N^+$ and is assumed to be independent of t and $\boldsymbol{\mu}$, $\mathbf{g}(\mathbf{y}(t), \dot{\mathbf{y}}(t); \boldsymbol{\mu})$ is the $\mathbb{R}^N$ vector representing the internal forces at time t and depends on $\mathbf{y}(t)$, $\dot{\mathbf{y}}(t)$ and $\boldsymbol{\mu}$, and finally, $\mathbf{f}(t; \boldsymbol{\mu})$ is the $\mathbb{R}^N$ vector of external forces at time t and depends on $\boldsymbol{\mu}$.

In Eq. (2.2), $[B]$ is a given matrix in $\mathbb{M}_{N,N_{\text{CD}}}$. It defines the N_{CD} constraints on $\mathbf{y}$, is assumed to be independent of t and $\boldsymbol{\mu}$ and to satisfy $[B]^T[B] = [I_{N_{\text{CD}}}]$.

The $\mathbb{R}^N$-valued solution $\{\mathbf{y}(t; \boldsymbol{\mu}), t \in [t_0, T]\}$ of Eqs. (2.1) to (2.2) depends on $\boldsymbol{\mu}$. The QoI (system observation) at time t is denoted by the vector $\mathbf{o}(t; \boldsymbol{\mu}) = (o_1(t; \boldsymbol{\mu}), \ldots, o_{m_\mathbf{o}}(t; \boldsymbol{\mu}))$ with values in $\mathbb{R}^{m_\mathbf{o}}$. This vector depends on $\boldsymbol{\mu}$ and is here written as

$$\mathbf{o}(t; \boldsymbol{\mu}) = \mathbf{h}(\mathbf{y}(t; \boldsymbol{\mu}), \dot{\mathbf{y}}(t; \boldsymbol{\mu}), \mathbf{f}(t; \boldsymbol{\mu}), t; \boldsymbol{\mu}) \quad , \quad t \in [t_0, T] \,, \tag{2.3}$$

where $\mathbf{h}$ is a given mapping.

2.2. Construction of a $\boldsymbol{\mu}$-parametric nonlinear reduced-order model

Let $[V] \in \mathbb{M}_{N,n}$ be a global ROB (independent of $\boldsymbol{\mu}$) of dimension $n << N$ constructed for approximating the solution $\{\mathbf{y}(t; \boldsymbol{\mu}), t \in [t_0, T]\}$ for all $\boldsymbol{\mu} \in \mathcal{C}_{\boldsymbol{\mu}}$. From (2.2), it follows that $[V]$ must satisfy the constraint equation

$$[B]^T[V] = [0_{N_{\text{CD}},n}] \,. \tag{2.4}$$

Typically, $[V]$ satisfies also by construction the orthonormality condition [13, 14]

$$[V]^T[M]\,[V] = [I_n] \,. \tag{2.5}$$

The Galerkin projection of the HDM represented by (2.1) on such a ROB leads to the $\boldsymbol{\mu}$-parametric nonlinear ROM

$$\mathbf{y}^{(n)}(t) = [V]\,\mathbf{q}(t) \quad , \quad t \in [t_0, T] \,, \tag{2.6}$$

$$\ddot{\mathbf{q}}(t) + [V]^T \mathbf{g}([V]\,\mathbf{q}(t), [V]\,\dot{\mathbf{q}}(t); \boldsymbol{\mu}) = [V]^T \mathbf{f}(t; \boldsymbol{\mu}) \quad , \quad t \in]t_0, T] \,, \tag{2.7}$$

with the initial conditions

$$\mathbf{q}(t_0) = [V]^T[M]\,\mathbf{y}_0 \quad , \quad \dot{\mathbf{q}}(t_0) = [V]^T[M]\,\mathbf{y}_1 \,, \tag{2.8}$$

where the $\mathbb{R}^n$-valued solution $\{\mathbf{q}(t; \boldsymbol{\mu}), t \in [t_0, T]\}$ of Eqs. (2.7) and (2.8) depends on $\boldsymbol{\mu}$ — and is often referred to as the vector of generalized coordinates — and $\{\mathbf{y}^{(n)}(t; \boldsymbol{\mu}), t \in [t_0, T]\}$ is the n-order approximation of $\{\mathbf{y}(t; \boldsymbol{\mu}), t \in [t_0, T]\}$. The corresponding approximation $\mathbf{o}^{(n)}$ of the QoI $\mathbf{o}$ is given by

$$\mathbf{o}^{(n)}(t; \boldsymbol{\mu}) = \mathbf{h}(\mathbf{y}^{(n)}(t; \boldsymbol{\mu}), \dot{\mathbf{y}}^{(n)}(t; \boldsymbol{\mu}), \mathbf{f}(t; \boldsymbol{\mu}), t; \boldsymbol{\mu}) \quad , \quad t \in [t_0, T] \quad , \quad \boldsymbol{\mu} \in \mathcal{C}_{\boldsymbol{\mu}} \,. \tag{2.9}$$

For a given n, the prediction error due to the use of the $\boldsymbol{\mu}$-parametric nonlinear ROM instead of the $\boldsymbol{\mu}$-parametric nonlinear HDM could, *a priori*, be estimated (in the L^2 sense) by

$$\varepsilon(n) = \int_{\mathcal{C}_{\boldsymbol{\mu}}} \int_{t_0}^{T} \|\mathbf{o}(t;\boldsymbol{\mu}) - \mathbf{o}^{(n)}(t;\boldsymbol{\mu})\|^2 \, dt \, d\boldsymbol{\mu} \, . \tag{2.10}$$

However, such an estimation entails the computation of the high-dimensional solution of the problem for a very large number of sampled values of the vector-valued parameter $\boldsymbol{\mu}$ in $\mathcal{C}_{\boldsymbol{\mu}}$ and therefore can be cost-prohibitive. For this reason, Eq. (2.10) is replaced in practice by

$$\varepsilon(n) = \sum_{i=1}^{m_\mu} \int_{t_0}^{T} \|\mathbf{o}(t;\boldsymbol{\mu}_i) - \mathbf{o}^{(n)}(t;\boldsymbol{\mu}_i)\|^2 \, dt \, ,$$

where $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}$ are the sampling points used for constructing the ROB $[V]$.

2.3. *Construction of a stochastic reduced-order model*

As already stated in the abstract, the main idea contributed by this paper includes substituting the deterministic ROB $[V]$ with a stochastic counterpart $[\mathbf{W}]$. In view of the properties of $[V]$, including the constraint (2.4) and orthonormality condition (2.5), this SROB must verify the following properties:

- $[\mathbf{W}]$ is global and therefore independent of $\boldsymbol{\mu}$, because $[V]$ is global and independent of $\boldsymbol{\mu}$.
- $[\mathbf{W}]$ is a random matrix with values in $\mathbb{M}_{N,n}$.
- The support of its probability distribution (constructed using the Maximum Entropy principle of Information Theory) is the subset of $\mathbb{M}_{N,n}$ corresponding to the constraints

  $$[B]^T[\mathbf{W}] = [0_{N_{\mathrm{cb}},n}] \tag{2.11}$$

  and

  $$[\mathbf{W}]^T[M]\,[\mathbf{W}] = [I_n] \tag{2.12}$$

  almost surely.
- The probability distribution of $[\mathbf{W}]$ depends on a vector-valued hyperparameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{m_{\boldsymbol{\alpha}}})$ belonging to a subset $\mathcal{C}_{\boldsymbol{\alpha}}$ of $\mathbb{R}^{m_{\boldsymbol{\alpha}}}$ where the dimension $m_{\boldsymbol{\alpha}}$ is chosen sufficiently small so that the statistical inverse problem for identifying the hyperparameter vector $\boldsymbol{\alpha}$ is computationally feasible.

The construction of such an SROB is described in Section 3.

Next, the $\boldsymbol{\mu}$-parametric, nonlinear, projection-based SROM associated with the $\boldsymbol{\mu}$-parametric, nonlinear, projection-based ROM described above is deduced from Eqs. (2.6) to (2.9) by substituting $[V]$ with the random matrix $[\mathbf{W}]$. Consequently, $\mathbf{y}^{(n)}$, $\mathbf{q}$, and $\mathbf{o}^{(n)}$ become the stochastic processes $\mathbf{Y}^{(n)}$, $\mathbf{Q}$, and $\mathbf{O}^{(n)}$, and the SROM can be written as

$$\mathbf{Y}^{(n)}(t) = [\mathbf{W}]\,\mathbf{Q}(t) \quad , \quad t \in [t_0, T] \, , \tag{2.13}$$

$$\ddot{\mathbf{Q}}(t) + [\mathbf{W}]^T \mathbf{g}([\mathbf{W}]\,\mathbf{Q}(t), [\mathbf{W}]\,\dot{\mathbf{Q}}(t); \boldsymbol{\mu}) = [\mathbf{W}]^T \mathbf{f}(t;\boldsymbol{\mu}) \quad , \quad t \in]t_0, T] \, , \tag{2.14}$$

with the initial conditions

$$\mathbf{Q}(t_0) = [\mathbf{W}]^T[M]\,\mathbf{y}_0 \quad , \quad \dot{\mathbf{Q}}(t_0) = [\mathbf{W}]^T[M]\,\mathbf{y}_1 \, . \tag{2.15}$$

The $\mathbb{R}^n$-valued stochastic solution $\{\mathbf{Q}(t;\boldsymbol{\mu},\boldsymbol{\alpha}), t \in [t_0,T]\}$ of Eqs. (2.14) and (2.15) depends on $\boldsymbol{\mu} \in \mathcal{C}_{\boldsymbol{\mu}}$ and $\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\alpha}}$. The stochastic process $\{\mathbf{Y}^{(n)}(t;\boldsymbol{\mu},\boldsymbol{\alpha}), t \in [t_0,T]\}$ is the n-order approximation of the stochastic process $\{\mathbf{Y}(t;\boldsymbol{\mu}), t \in [t_0,T]\}$. The corresponding approximation $\{\mathbf{O}^{(n)}(t;\boldsymbol{\mu},\boldsymbol{\alpha}), t \in [t_0,T]\}$ of the random QoI $\{\mathbf{O}(t;\boldsymbol{\mu}), t \in [t_0,T]\}$ is given, for all $t \in [t_0,T]$, $\boldsymbol{\mu} \in \mathcal{C}_{\boldsymbol{\mu}}$ and $\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\alpha}}$, by

$$\mathbf{O}^{(n)}(t;\boldsymbol{\mu},\boldsymbol{\alpha}) = \mathbf{h}(\mathbf{Y}^{(n)}(t;\boldsymbol{\mu},\boldsymbol{\alpha}), \dot{\mathbf{Y}}^{(n)}(t;\boldsymbol{\mu},\boldsymbol{\alpha}), \mathbf{f}(t;\boldsymbol{\mu}), t; \boldsymbol{\mu}) \, . \tag{2.16}$$

2.4. *Identification of the hyperparameter vector of the probability distribution of the stochastic reduced-order basis*

The identification of the hyperparameter vector $\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\alpha}} \subset \mathbb{R}^{m_{\boldsymbol{\alpha}}}$ can be performed using the maximum likelihood method, or a nonlinear Least-Squares (LS) method for the QoI. For example, a nonlinear LS method can be formulated as follows for both types of modeling errors highlighted in Section 1.5.

Let $J(\boldsymbol{\alpha})$ be the cost function defined on $\mathcal{C}_{\boldsymbol{\alpha}}$ by

$$J(\boldsymbol{\alpha}) = w_J\, J_{\text{mean}}\,(\boldsymbol{\alpha}) + (1 - w_J)\, J_{\text{std}}\,(\boldsymbol{\alpha})\,, \tag{2.17}$$

where w_J is a weight satisfying $0 \leq w_J \leq 1$, and $J_{\text{mean}}\,(\boldsymbol{\alpha})$ and $J_{\text{std}}\,(\boldsymbol{\alpha})$ allow for controlling the identification of $\boldsymbol{\alpha}$ with respect to the mean value and statistical fluctuations, respectively. The latter quantities are defined here as

$$J_{\text{mean}}\,(\boldsymbol{\alpha}) = \frac{1}{c_{\text{mean}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu})} \sum_{i=1}^{m_\mu} \int_{t_0}^{T} \|\mathbf{o}^{\text{ref}}(t; \boldsymbol{\mu}_i) - E\{\mathbf{O}^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha})\}\|^2\, dt\,, \tag{2.18}$$

$$J_{\text{std}}\,(\boldsymbol{\alpha}) = \frac{1}{c_{\text{std}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu})} \sum_{i=1}^{m_\mu} \int_{t_0}^{T} \|\mathbf{v}^{(\text{ref},n)}(t; \boldsymbol{\mu}_i) - \mathbf{v}^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha})\|^2\, dt\,, \tag{2.19}$$

where the positive constants $c_{\text{mean}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu})$ and $c_{\text{std}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu})$ are given by

$$c_{\text{mean}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}) = \sum_{i=1}^{m_\mu} \int_{t_0}^{T} \|\mathbf{o}^{\text{ref}}(t; \boldsymbol{\mu}_i)\|^2\, dt\,, \tag{2.20}$$

$$c_{\text{std}}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}) = \sum_{i=1}^{m_\mu} \int_{t_0}^{T} \|\mathbf{v}^{(\text{ref},n)}(t; \boldsymbol{\mu}_i)\|^2\, dt\,,$$

and $\mathbf{v}^{(\text{ref},n)}(t; \boldsymbol{\mu}_i) = (v_1^{(\text{ref},n)}(t; \boldsymbol{\mu}_i), \ldots, v_{m_{\mathbf{o}}}^{(\text{ref},n)}(t; \boldsymbol{\mu}_i))$ is such that,

$$v_j^{(\text{ref},n)}(t; \boldsymbol{\mu}_i) = \gamma\, |o_j^{\text{ref}}(t; \boldsymbol{\mu}_i) - o_j^{(n)}(t; \boldsymbol{\mu}_i)| \quad , \quad j = 1, \ldots m_{\mathbf{o}}\,, \tag{2.21}$$

where $\gamma > 0$ allows to control the amplitude of the target related to the statistical fluctuations. In Eq. (2.19), $\mathbf{v}^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha}) = (v_1^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha}), \ldots, v_{m_{\mathbf{o}}}^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha}))$ is such that

$$v_j^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha}) = \{\, E\{O_j^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha})^2\} - (E\{O_j^{(n)}(t; \boldsymbol{\mu}_i, \boldsymbol{\alpha})\})^2\} \,\}^{1/2} \quad , \quad j = 1, \ldots m_{\mathbf{o}}\,.$$

In Eqs. (2.18), (2.20) and (2.21), $\mathbf{o}^{\text{ref}}$ is defined as a function of the type of the modeling errors that are taken into account:

$\diamond$ If only the errors due to model reduction are taken into account, then $\mathbf{o}^{\text{ref}} = \mathbf{o}$.

$\diamond$ If the errors due to model reduction and those introduced in the construction of the underlying $\boldsymbol{\mu}$-parametric nonlinear HDM are taken into account, then $\mathbf{o}^{\text{ref}} = \mathbf{o}^{\text{exp}}$, where $\mathbf{o}^{\text{exp}}$ is based on experimental data.

It follows that in the cost function defined above:

$\diamond$ $\mathbf{o}^{\text{ref}}$ appears as the target for the mean value with a weight w_J.

$\diamond$ $\mathbf{v}^{(\text{ref},n)}$ appears as the target for the standard deviation with a weight $1 - w_J$.

Hence, the identification of the hyperparameter vector $\boldsymbol{\alpha}$ consists in calculating $\boldsymbol{\alpha}^{\text{opt}}$ such that

$$\boldsymbol{\alpha}^{\text{opt}} = \min_{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\alpha}}} J(\boldsymbol{\alpha})\,.$$

2.5. Justification

At this point, the reader may ask why is a random basis of the admissible set useful? The answer to this question has three parts:

- Any deterministic global ROB $[V]$ of dimension $n = N$ constitutes a basis of the admissible vector space for the solutions of the $\boldsymbol{\mu}$-parametric HDM, for all $\boldsymbol{\mu}$ belonging to $\mathcal{C}_{\boldsymbol{\mu}}$. On the other hand, for a fixed dimension $n < N$ and a given sampling $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}$ in $\mathcal{C}_{\boldsymbol{\mu}}$, the solution of an instance of the $\boldsymbol{\mu}$-parametric HDM problem formulated using an unsampled parameter vector $\boldsymbol{\mu}^\star$ cannot be necessarily approximated in the subspace spanned by $[V]$ with a specified level of accuracy. Using the proposed nonparametric probabilistic approach however, the variability of $[V]$ can be captured for the following reason. Every ROB sample generated by the constructed SROB constitutes a family of algebraically independent vectors that verifies the boundary conditions of the problem of interest, satisfies the regularity required for solving the boundary value problem represented by the HDM, and represents a fluctuation around $[V]$ whose magnitude is controlled by the hyperparameters of the SROB, but can be large enough to capture any variability of $[V]$ while satisfying all of the required constraints. Therefore, the proposed nonparametric probabilistic approach for taking into account model form uncertainties can be interpreted as a stochastic-based method for extracting fundamental information or knowledge from test or HDM data that is not captured by a deterministic HDM or ROM. Using this approach, one essentially parameterizes the approximation basis in order to capture the variabilities instead of (artificially) parameterizing the governing equations.
- In the deterministic setting, it is well known that a good approximation of order n, $\mathbf{y}^{(n)}$, of the solution $\mathbf{y}$ of a problem of interest delivered by an HDM of dimension N can be computed using a Galerkin projection method — for example, a projection-based ROM of dimension n — independently of the explicit choice of the admissible vector basis. For example, this statement is trivial for $n = N$. Similarly, in the stochastic setting, a family of stochastic approximations $\mathbf{Y}^{(n)}$ that converges almost surely towards the deterministic solution $\mathbf{y}$ based on an HDM can be constructed using an SROM. At convergence, $\mathbf{Y}^{(n)}$ is independent of the choice of the stochastic basis. However, for $n << N$, and taking into account the fact that the deterministic ROB $[V]$ is constructed for $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}$ in $\mathcal{C}_{\boldsymbol{\mu}}$, for any unsampled value of $\boldsymbol{\mu}$ in $\mathcal{C}_{\boldsymbol{\mu}}$, the approximation $\mathbf{y}^{(n)}$ delivered by the associated ROM is tainted by both sampling- and truncation (or reduction)-induced errors. In this case, for a fixed value of n for which convergence is not reached, the value of the approximation error depends on the choice of the admissible basis. Subsequently, when the ROB $[V]$ is substituted with the SROB $[\mathbf{W}]$, a stochastic family of approximations $\mathbf{Y}^{(n)}$ is generated using the SROM. Hence, the idea is to adapt the statistical fluctuations of $\mathbf{Y}^{(n)}$ for representing not only the sampling and model reduction related errors, but also the modeling errors introduced during the construction of the $\boldsymbol{\mu}$-parametric nonlinear HDM itself.
- The proposed nonparametric probabilistic approach underlying the usage of a random basis of the admissible set can account for experimental variability in two different ways: qualitatively, by the probabilistic nature of the SROM that generates the statistical fluctuations of the QoI; and quantitatively, by choosing the experimental data as the target for the QoI in the identification procedure of the hyperparameters of the SROB.

3. CONSTRUCTION OF A STOCHASTIC REDUCED-ORDER BASIS

From Eq. (2.12), it follows that the $N \times n$ SROB $[\mathbf{W}]$ must satisfy an orthonormality condition. Hence, it must be constructed on a compact Stiefel manifold denoted here by $\mathbb{S}_{N,n}$. From Eq. (2.11), it follows that this SROB must also satisfy an additional constraint equation. Therefore, more specifically, it must be constructed on a subset of $\mathbb{S}_{N,n}$. For this reason, a brief summary of known results concerning the parameterization of $\mathbb{S}_{N,n}$ is given below. Then, the complete construction of the SROB $[\mathbf{W}]$ and the analysis of its mathematical properties are presented in Sections 3.3 to 3.5.

3.1. The compact Stiefel manifold: tangent vector space and parameterizations

Let n and N be two integers satisfying $1 < n << N$, and let $[M] \in \mathbb{M}_N^+$ denote an $(N \times N)$ real, SPD matrix. The set $\mathbb{M}_{N,n}$ of all $(N \times n)$ real matrices forms an Euclidean space equipped with the inner product and associated norm defined by

$$\ll [V_1]\,, [V_2] \gg_M = \text{tr}\{[V_1]^T\,[M]\,[V_2]\} \quad , \quad \|V\|_M = \{\text{tr}\{[V]^T\,[M]\,[V]\}\}^{1/2}\,.$$

The compact Stiefel manifold $\mathbb{S}_{N,n}$ is defined as (for example, see [65])

$$\mathbb{S}_{N,n} = \{\ [V] \in \mathbb{M}_{N,n}\ ,\ [V]^T[M]\,[V] = [I_n]\ \} \subset \mathbb{M}_{N,n}\,. \tag{3.1}$$

Its dimension is $\nu_{\mathbb{S}} = Nn - n(n+1)/2$, which can also be written as $\nu_{\mathbb{S}} = n(n-1)/2 + n(N-n)$.

From $[V]^T[M]\,[V] = [I_n]$, it follows that $[V]^T[M]\,[dV] + [dV]^T[M]\,[V] = [0_{n,n}]$, where $[dV]$ denotes the 1-differential form of $[V]$. Therefore, the tangent *vector space* to $\mathbb{S}_{N,n}$ at the point $[V] \in \mathbb{S}_{N,n}$, $T_V\mathbb{S}_{N,n}$, is defined by

$$T_V\mathbb{S}_{N,n} = \{\ [Z] \in \mathbb{M}_{N,n} : [V]^T[M]\,[Z] + [Z]^T[M]\,[V] = [0_{n,n}]\ \}\,. \tag{3.2}$$

In the context of this work, the tangent vector space $T_V\mathbb{S}_{N,n}$ plays an important role for two reasons:

- It provides a mean for parameterizing $\mathbb{S}_{N,n}$.
- An explicit mapping can be constructed between this tangent vector space and $\mathbb{S}_{N,n}$. Hence, if for constructing the random matrix [**W**] the fluctuations around the deterministic ROB $[V]$ are generated in this tangent vector space, two important facts become noteworthy: (1) the resulting perturbed ROB remains in this vector space, and therefore (2) it can be mapped onto $\mathbb{S}_{N,n}$ to obtain a matrix [**W**] that satisfies the constraint equations (2.11) and (2.12).

From the definition (3.2), it follows that any element $[Z]$ of $T_V\mathbb{S}_{N,n}$ can be written as

$$[Z] = [V]\,[a] + [V_\perp]\,[b]\,, \tag{3.3}$$

where $[V_\perp]$ denotes an element of $\mathbb{M}_{N,N-n}$ satisfying

$$[V]^T[M]\,[V_\perp] = [0_{n,N-n}]\,,$$

$[a]$ is a skew-symmetric matrix in $\mathbb{M}_n^{\text{SS}}$ ($[a]^T = -[a]$), and $[b]$ is any matrix in $\mathbb{M}_{N-n,n}$. Eq. (3.3) constitutes a parameterization of the tangent vector space $T_V\mathbb{S}_{N,n}$ at a given point $[V]$ in $\mathbb{S}_{N,n}$. Specifically, $n(n-1)/2$ entries of $[a]$ and $n(N-n)$ entries of $[b]$ parameterize $T_V\mathbb{S}_{N,n}$ at the point $[V]$. Therefore, the dimension of $T_V\mathbb{S}_{N,n}$ is $\nu_{T_V\mathbb{S}_{N,n}} = n(n-1)/2 + n(N-n)$ — that is, $\nu_{T_V\mathbb{S}_{N,n}} = \nu_{\mathbb{S}}$.

The parameterization (3.3) of $T_V\mathbb{S}_{N,n}$ requires the construction of a large-scale matrix $[V_\perp]$ that belongs to $\mathbb{M}_{N,N-n}$. For large values of N, this is a major drawback. In this case, the following alternative parameterization is preferred.

Let $[A]$ denote an arbitrary matrix in $\mathbb{M}_{N,n}$. Its projection onto $T_V\mathbb{S}_{N,n}$, $\text{Proj}_{T_V\mathbb{S}_{N,n}}([A])$, can be written as

$$[Z] = \text{Proj}_{T_V\mathbb{S}_{N,n}}([A]) := [A] - [V]\,[D] \quad , \quad [D] = ([V]^T[M]\,[A] + [A]^T[M]\,[V])/2\,, \tag{3.4}$$

where $[D] \in \mathbb{M}_n^S$ is an $(n \times n)$ symmetric matrix. The above result can also be written as $[Z] = [V]\,[a] + ([I_N] - [V]\,[V]^T[M])[A]$, where $[a] = ([V]^T[M]\,[A] - [A]^T[M]\,[V])/2$ belongs to $\mathbb{M}_n^{\text{SS}}$. In this case, $[V]^T[M]\,[Z] = [a]$, and given that $\text{tr}[a] = 0$, $\ll [V]\,, [Z] \gg = 0$.

REMARK 1. For large values of N, the dimension of the parameterization (3.4) of $T_V\mathbb{S}_{N,n}$, Nn, is much smaller than that of its counterpart (3.3). Still, the dimension of this parameterization

is larger than the dimension of $T_V\mathbb{S}_{N,n}$ itself, $\nu_{\mathbb{S}} = Nn - n(n+1)/2$. In particular, note that $Nn = \nu_{\mathbb{S}} + n(n+1)/2$, which reveals that there are $n(n+1)/2$ unnecessary parameters in the representation (3.4) of $T_V\mathbb{S}_{N,n}$. These unnecessary parameters are associated with the normal space to $\mathbb{S}_{N,n}$ at a given point $[V]$ in $\mathbb{S}_{N,n}$, $N_V\mathbb{S}_{N,n}$. Indeed, from (3.4), it follows that $N_V\mathbb{S}_{N,n}$ consists of the set of matrices of the form $[A_S] = [V]\,[S]$, where $[S]$ is a symmetric matrix in $\mathbb{M}_n^S$ ($[Z] = \text{Proj}_{T_V\mathbb{S}_{N,n}}([A_S]) = [0_{N,n}]$). The dimension of this vector space is $n(n+1)/2$. Nevertheless, for large values of N, the parameterization of $T_V\mathbb{S}_{N,n}$ defined by (3.4) is computationally feasible for constructing the random matrix $[\mathbf{W}]$ associated with the stochastic modeling of $[V] \in \mathbb{S}_{N,n}$, whereas that defined by (3.3) is not computationally feasible for this purpose.

Now, let $[Z] \mapsto R_{s,V}([Z])$ denote the smooth mapping from $T_V\mathbb{S}_{N,n}$ to $\mathbb{S}_{N,n}$,

$$[W] = R_{s,V}([Z])\,,$$

that verifies

$$[V] = R_{s,V}([0_{N,n}])\,.$$

Then, for any $[Z]$ in $T_V\mathbb{S}_{N,n}$, $[W] = R_{s,V}([Z])$ satisfies $[W]^T[M]\,[W] = [I_n]$.

There are several possibilities for constructing the mapping $R_{s,V}$. For example, one can note the construction procedure based on the economy-size QR decomposition [67], which is computationally efficient for N very large and $n << N$. In this work, the following equally efficient procedure based on the adaptation of the polar decomposition [66] to Eq. (3.1) is used

$$[W] = R_{s,V}([Z]) := ([V] + s[Z])\,[H_s(Z)] \quad , \quad [Z] \in T_V\mathbb{S}_{N,n}\,. \tag{3.5}$$

In (3.5) above, $s \geq 0$ is a real number introduced for controlling the level of fluctuations of $[Z]$ in $T_V\mathbb{S}_{N,n}$ around $[V]$ in $\mathbb{S}_{N,n}$, and $[H_s(Z)]$ is the SPD matrix in $\mathbb{M}_n^+$ given by

$$[H_s(Z)] = ([I_n] + s^2[Z]^T[M]\,[Z])^{-1/2}\,. \tag{3.6}$$

The construction procedure defined by (3.5) and (3.6) verifies $R_{s,V}([0_{N,n}]) = [V]$. This means that when $[Z]$ undergoes small fluctuations around $[0_{N,n}]$, $[W] = R_{s,V}([Z])$ undergoes fluctuations around $[V]$ and their amplitude can be controlled by the parameter s for a given normalization of $[Z]$. Now, substituting the parameterization of $[Z]$ given in (3.4) into the definition of $[W] = R_{s,V}([Z])$ given by Eqs. (3.5) and (3.6) leads to the mapping $[A] \mapsto \mathcal{R}_{s,V}([A])$ from $\mathbb{M}_{N,n}$ into $\mathbb{S}_{N,n}$ defined for all $[A]$ in $\mathbb{M}_{N,n}$ by

$$[W] = \mathcal{R}_{s,V}([A]) := R_{s,V}([A] - [V]\,[D]) \quad , \quad [D] = ([V]^T[M]\,[A] + [A]^T[M]\,[V])/2\,. \tag{3.7}$$

In particular, using the above result, the reader can verify that

$$[W]^T[M]\,[W] = [I_n] \quad , \quad [V] = \mathcal{R}_{s,V}([0_{N,n}]) \in \mathbb{S}_{N,n}\,.$$

3.2. Parameterization of a subset of a compact Stiefel manifold associated with additional constraints

Consider now the case where $[V]$ belongs to the subset $\mathcal{S}_{N,n}$ of $\mathbb{S}_{N,n}$ defined by

$$\mathcal{S}_{N,n} = \{\ [V] \in \mathbb{M}_{N,n}\ ,\ [V]^T[M]\,[V] = [I_n]\ ,\ [B]^T[V] = [0_{N_{\text{CD}},n}]\ \} \subset \mathbb{S}_{N,n}\,, \tag{3.8}$$

where $0 < N_{\text{CD}} < N$ is the number of constraint equations, $[B]$ is a given matrix satisfying

$$[B] \in \mathbb{M}_{N,N_{\text{CD}}} \quad , \quad [B]^T\,[B] = [I_{N_{\text{CD}}}]\,, \tag{3.9}$$

and $[B]^T[V] = [0_{N_{\text{CD}},n}]$ represents a set of N_{CD} linear constraints expressed here in orthonormal form and associated, for example, with Dirichlet boundary conditions and/or kinematic constraints between some dofs of the computational model for which the ROB $[V]$ is constructed. Then, the

objective is to extend the parameterization $[W] = \mathcal{R}_{s,V}([A]) \in \mathbb{S}_{N,n}$ defined by Eqs. (3.7) and (3.1) to a parameterization $[W] = \mathcal{R}_{s,V}([A]) \in \mathcal{S}_{N,n} \subset \mathbb{S}_{N,n}$, so that the additional constraint equation $[B]^T[W] = [0_{N_{\text{CD}},n}]$ is satisfied. Such a parameterization, which must verify

$$[W]^T[M]\,[W] = [I_n] \quad , \quad [B]^T[W] = [0_{N_{\text{CD}},n}] \quad , \quad [V] = \mathcal{R}_{s,V}([0_{N,n}])\,, \tag{3.10}$$

is given by Eq. (3.7), where $[A]$ belongs in this case to the subset $\mathbb{M}^B_{N,n}$ of $\mathbb{M}_{N,n}$ defined by

$$\mathbb{M}^B_{N,n} = \{\ [A] \in \mathbb{M}_{N,n}\ ,\ [A] = ([I_N] - [B]\,[B]^T)\,[U]\ ,\ [U] \in \mathbb{M}_{N,n}\ ,\}\,. \tag{3.11}$$

Indeed, using (3.9), the reader can verify that for all $[A]$ in $\mathbb{M}^B_{N,n}$, $[B]^T[A] = [0_{N_{\text{CD}},n}]$. Using the polar representation of $R_{s,V}$ (see Eqs. (3.5) and (3.6)) with the parameterization of $[Z]$ defined by Eq. (3.4), and $[B]^T[V] = [0_{N_{\text{CD}},n}]$ (see Eq. (3.8)), the user can also verify that $[B]^T[W] = [0_{N_{\text{CD}},n}]$.

It is important to note that from a numerical point of view, the parameterization of the matrix $[A]$ introduced in (3.11) can be rewritten as $[A] = [U] - [B]\,\{[B]^T\,[U]\}$, where $[B]^T\,[U] \in \mathbb{M}_{N_{\text{CD}},n}$. Therefore, the $(N \times N)$ matrix $[B]\,[B]^T$ does not need to be assembled in practice. Furthermore, if the constraint $[B]^T[W] = [0_{N_{\text{CD}},n}]$ does not apply, Eq. (3.11) is simply replaced by $[A] = [U]$, where $[U] \in \mathbb{M}_{N,n}$.

Next, given a global ROB $[V]$, the associated SROB $[\mathbf{W}]$ is constructed on $\mathbb{S}_{N,n}$ with the additional constraint equation $[B]^T[W] = [0_{N_{\text{CD}},n}]$. Using the notation introduced above, this means that the matrix $[A]$ in $\mathbb{M}^B_{N,n}$ is given by (3.11). If no additional constraint equation of the form $[B]^T[W] = [0_{N_{\text{CD}},n}]$ is specified, $[A] = [U]$. Hence, the stochastic modeling of the matrix $[U]$ described next will be the same for both cases of a ROB $[V]$ with or without additional constraint equation.

3.3. Construction of a stochastic reduced-order basis

As anticipated in Section 2.3, given a ROB $[V]$ in $\mathcal{S}_{N,n} \subset \mathbb{S}_{N,n}$ (and therefore satisfying $[B]^T[V] = [0_{N_{\text{CD}},n}]$), the corresponding SROB is constructed here as a random matrix $[\mathbf{W}]$ defined on a probability space $(\Theta, \mathcal{T}, \mathcal{P})$ with values in $\mathcal{S}_{N,n}$ — that is, verifying the conditions (2.12) and (2.11) — and possibly additional constraints that are defined later. The construction of the probability measure $P_{[\mathbf{W}]}$ of $[\mathbf{W}]$ on $\mathbb{M}_{N,n}$, for which the support is the manifold $\mathcal{S}_{N,n}$,

$$\text{supp}\, P_{[\mathbf{W}]} = \mathcal{S}_{N,n} \subset \mathbb{S}_{N,n} \subset \mathbb{M}_{N,n}\,,$$

requires the introduction of an adapted parameterization of $\mathcal{S}_{N,n}$ that addresses the difficulties induced by the support of the measure. To this effect, a three-step construction procedure for $[\mathbf{W}]$, in which the available information is gradually introduced, is presented below.

3.3.1. Step 1: parameterization The SROB is constructed using the parameterization of $[W]$ given in (3.7) and (3.11). The corresponding random matrix $[\mathbf{W}]$ can then be written as

$$[\mathbf{W}] = \mathcal{R}_{s,V}([\mathbf{A}]) = R_{s,V}([\mathbf{A}] - [V]\,[\mathbf{D}]) \quad , \quad [\mathbf{D}] = ([V]^T[M]\,[\mathbf{A}] + [\mathbf{A}]^T[M]\,[V])/2\,, \tag{3.12}$$

where $[\mathbf{A}]$ is a random matrix defined on $(\Theta, \mathcal{T}, \mathcal{P})$, with values in the subset $\mathbb{M}^B_{N,n}$ of $\mathbb{M}_{N,n}$. Hence, $[\mathbf{A}]$ can be written as

$$[\mathbf{A}] = ([I_N] - [B]\,[B]^T)\,[\mathbf{U}] = [\mathbf{U}] - [B]\,\{[B]^T\,[\mathbf{U}]\}\,, \tag{3.13}$$

where the random matrix $[\mathbf{U}] = [\mathbf{U}^1 \dots \mathbf{U}^n]$ is defined on $(\Theta, \mathcal{T}, \mathcal{P})$, with values in $\mathbb{M}_{N,n}$. The columns of $[\mathbf{U}]$ are n random vectors $\mathbf{U}^1, \dots, \mathbf{U}^n$ with values in $\mathbb{R}^N$. The deterministic mapping $R_{s,V}$ (from the tangent vector space $T_V\mathbb{S}_{N,n}$ of $\mathbb{S}_{N,n}$ at point $[V]$ to $\mathbb{S}_{N,n}$) is defined by Eqs. (3.5) and (3.6). Note that $[\mathbf{D}]$ is a random matrix with values in $\mathbb{M}^S_n$.

3.3.2. Step 2: defining the available information Taking into account Eqs. (3.12) and (3.13), the stochastic model of the random matrix $[\mathbf{W}]$ and its generator are completely defined by their counterparts for the random matrix $[\mathbf{U}]$.

A main requirement for the construction of the SROB is that the statistical fluctuations of the random matrix $[\mathbf{W}]$ be around the deterministic matrix $[V]$. Taking into account that $[\mathbf{W}] = [V]$ for $[\mathbf{A}] = [0_{N,n}]$ (due to the third of Eqs. (3.10)), the random matrix $[\mathbf{A}]$ must be a centered random variable, which is satisfied if $[\mathbf{U}]$ is also a centered random matrix (due to Eq. (3.13)). Hence, one must have

$$E\{[\mathbf{U}]\} = [0_{N,n}]\,. \tag{3.14}$$

Furthermore, in order to minimize the number of hyperparameters in the stochastic model of the centered random matrix $[\mathbf{U}]$, the $Nn(Nn+1)/2$ components of the fourth-order symmetric covariance tensor $\{\mathbb{c}_{jkj'k'}\}_{jkj'k'}$ of the $\mathbb{M}_{N,n}$-valued random matrix $[\mathbf{U}]$ are not kept, because N can be very large. Therefore, the following reduced parameterization with $1 + n(n+1)/2$ hyperparameters is chosen for tensor $\mathbb{c}$. Specifically, for all j and j' in $\{1,\ldots,N\}$, and for all k and k' in $\{1,\ldots,n\}$,

$$\mathbb{c}_{jkj'k'} = E\{\mathbf{U}_{jk}\mathbf{U}_{j'k'}\} = [C_N(\beta)]_{jj'}[c_n]_{kk'}\,, \tag{3.15}$$

where:

- $[C_N(\beta)] \in \mathbb{M}_N^+$ is a type-covariance matrix depending on a hyperparameter β such that $0 < \beta_d \leq \beta \leq \beta_u < +\infty$, and is constructed in Appendix D (see Eq. (D.57)). This covariance matrix allows the introduction of a correlation between the components $U_1^k,\ldots,U_N^k$ of each random vector $\mathbf{U}^k$ so that $U_j^k = [\mathbf{U}]_{jk}$.
- $[c_n] \in \mathbb{M}_n^+$ is a type-covariance matrix: therefore, there exists an upper triangular matrix $[\sigma]$ belonging to $\mathbb{M}_n^{\mathrm{u}}$ so that (Cholesky's factorization)

$$[c_n] = [\sigma]^T[\sigma]\,. \tag{3.16}$$

  Specifically, the matrix $[c_n]$ allows to describe the correlation between the random vectors $\mathbf{U}^1,\ldots,\mathbf{U}^n$.

Using Eqs. (3.14), (3.15), and (3.16), the second-order $\mathbb{M}_{N,n}$-valued random matrix $[\mathbf{U}]$ can be rewritten as $[\mathbf{U}] = [\mathbf{G}]\,[\sigma]$, where $[\mathbf{G}]$ is a second-order centered $\mathbb{M}_{N,n}$-valued random matrix defined on the probability space $(\Theta,\mathcal{T},\mathcal{P})$, such that for all j and j' in $\{1,\ldots,N\}$, and for all k and k' in $\{1,\ldots,n\}$,

$$E\{\mathbf{G}_{jk}\} = 0 \quad , \quad E\{\mathbf{G}_{jk}\mathbf{G}_{j'k'}\} = [C_N(\beta)]_{jj'}\,\delta_{kk'}\,. \tag{3.17}$$

It follows that

$$E\{[\mathbf{G}]\,[\mathbf{G}]^T\} = n\,[C_N(\beta)] \quad , \quad E\{[\mathbf{G}]^T[\mathbf{G}]\} = (\mathrm{tr}[C_N(\beta)])\,[I_n]\,.$$

Therefore, the random matrix $[\mathbf{U}]$ is parameterized as

$$[\mathbf{U}] = [\mathbf{G}]\,[\sigma]\,, \tag{3.18}$$

and is such that

$$E\{[\mathbf{U}]\,[\mathbf{U}]^T\} = (\mathrm{tr}[c_n])\,[C_N(\beta)] \quad , \quad E\{[\mathbf{U}]^T[\mathbf{U}]\} = (\mathrm{tr}[C_N(\beta)])\,[c_n]\,.$$

3.3.3. Step 3: stochastic model for $[\boldsymbol{W}]$ In this step, the stochastic model and generator of independent realizations of the SROB $[\mathbf{W}]$ are constructed.

To this effect, it is first noted that the stochastic model of the second-order, centered, random matrix $[\mathbf{G}]$ with values in $\mathbb{M}_{N,n}$ is defined by Eq. (D.55) in Appendix D, and its generator of independent realizations is described in Section D.5 of this Appendix. The stochastic model is constructed so that it exhibits only a scalar hyperparameter β, and does not require the explicit construction of the large matrix $[C_N(\beta)]$ with $N(N+1)/2$ entries.

On the other hand, the stochastic representation of the second-order, uncentered, random matrix $[\mathbf{W}]$ with values in the manifold $\mathcal{S}_{N,n} \subset \mathbb{S}_{N,n} \subset \mathbb{M}_{N,n}$ is defined by Eqs. (3.12), (3.13), and (3.18),

which can be rewritten (using Eqs. (3.5) and (3.6)) as follows

$$[\mathbf{W}] = R_{s,V}([\mathbf{Z}]) = ([V] + s\,[\mathbf{Z}])\,[H_s(\mathbf{Z})]\,, \tag{3.19}$$

$$[H_s(\mathbf{Z})] = ([I_n] + s^2\,[\mathbf{Z}]^T[M]\,[\mathbf{Z}])^{-1/2}\,, \tag{3.20}$$

$$[\mathbf{Z}] = [\mathbf{A}] - [V]\,[\mathbf{D}]\,, \tag{3.21}$$

$$[\mathbf{D}] = ([V]^T[M]\,[\mathbf{A}] + [\mathbf{A}]^T[M]\,[V])/2\,, \tag{3.22}$$

$$[\mathbf{A}] = [\mathbf{U}] - [B]\,\{[B]^T\,[\mathbf{U}]\}\,, \tag{3.23}$$

$$[\mathbf{U}] = [\mathbf{G}(\beta)]\,[\sigma]\,, \tag{3.24}$$

where:

- $[\mathbf{G}(\beta)]$ is the second-order, centered, random matrix with values in $\mathbb{M}_{N,n}$ defined by Eq. (D.55) in Appendix D, and for which the covariance tensor is defined by Eq. (3.17).
- $[\sigma]$ is a given upper triangular matrix in $\mathbb{M}_n^{\mathrm{u}}$ (positive diagonal entries).
- $[B]$ is a given matrix in $\mathbb{M}_{N,N_{\mathrm{cd}}}$ verifying $[B]^T\,[B] = [I_{N_{\mathrm{cd}}}]$.
- $[V]$ is a given matrix in $\mathcal{S}_{N,n}$.

For $[V]$ fixed in $\mathcal{S}_{N,n}$, the $2 + n(n+1)/2$ *hyperparameters* of the stochastic model of the random matrix $[\mathbf{W}]$ with values in $\mathcal{S}_{N,n}$ are:

- The deterministic real parameter s, which is such that $\varepsilon_0 \leq s \leq 1$, where ε_0 is such that $0 \leq \varepsilon_0 < 1$ and is given (if $s = 0$, then $[\mathbf{W}] = [V]$ is deterministic and there are no statistical fluctuations).
- The deterministic real parameter β, which is such that $0 < \beta_d \leq \beta \leq \beta_u < +\infty$, where β_d and β_u are given.
- The upper triangular matrix $[\sigma]$ in $\mathbb{M}_n^{\mathrm{u}}$ (positive diagonal entries), which is parameterized by $n(n+1)/2$ parameters, and such that its diagonal entries satisfy the constraints $\varepsilon_0 \leq [\sigma]_{11}, \ldots, [\sigma]_{nn} \leq \sigma_u < +\infty$, where σ_u is given.

Thus, the hyperparameter vector is $\boldsymbol{\alpha} = (s, \beta, \{[\sigma]_{kk'}, 1 \leq k \leq k' \leq n\})$ with length $m_{\boldsymbol{\alpha}} = 2 + n(n+1)/2$. It belongs to the admissible set $\mathcal{C}_{\boldsymbol{\alpha}}$ defined by

$$\mathcal{C}_{\boldsymbol{\alpha}} = \{s \in [\varepsilon_0\,, 1]\,,\ \beta \in [\beta_d, \beta_u]\,,\ \varepsilon_0 \leq [\sigma]_{11}, \ldots, [\sigma]_{nn} \leq \sigma_u\,,\ [\sigma]_{kk'} \in \mathbb{R}\,, k < k'\}\,. \tag{3.25}$$

For the sake of simplicity, the notation (s, β, σ) is used for $\boldsymbol{\alpha}$ in the remainder of this paper.

REMARK 2. A good choice for the number of hyperparameters $m_{\boldsymbol{\alpha}}$ is problem dependent. For this reason, setting $m_{\boldsymbol{\alpha}}$ requires practical experience. To this effect, Section 3.3.2 provides some guidelines. For all *practical* purposes however, $m_{\boldsymbol{\alpha}}$ should be chosen as large as the computational model can afford from the computational complexity viewpoint. It should also take into account the number of nonzero elements in the matrix $[\sigma]$ whose sparsity can be controlled. Its smallest possible value corresponds to the smallest possible value of n — that is, $n = 1$ — and therefore is $m_{\boldsymbol{\alpha}} = 3$; in this case, $[\sigma]$ is the identity matrix multiplied by a scalar hyperparameter.

The generator of independent realizations of the random matrix $[\mathbf{W}]$ with values in $\mathcal{S}_{N,n}$ is directly given by Eqs. (3.19) to (3.24) and Eq. (D.55). For a fixed value of the hyperparameters s, β, and $[\sigma]$, and for $\theta \in \Theta$, the realization $[\mathbf{W}(\theta)]$ of $[\mathbf{W}]$ is computed as follows:

- Compute $[\mathbf{G}(\theta; \beta)]$ the generator described in Section D.5 of Appendix D.

- Then compute

$$
\begin{aligned}
[\mathbf{U}(\theta)] &= [\mathbf{G}(\theta;\beta)]\,[\sigma]\,,\\
[\mathbf{A}(\theta)] &= [\mathbf{U}(\theta)] - [B]\,\{[B]^T\,[\mathbf{U}(\theta)]\}\,,\\
[\mathbf{D}(\theta)] &= ([V]^T[M]\,[\mathbf{A}(\theta)] + [\mathbf{A}(\theta)]^T[M]\,[V])/2\,,\\
[\mathbf{Z}(\theta)] &= [\mathbf{A}(\theta)] - [V]\,[\mathbf{D}(\theta)]\,,\\
[H_s(\mathbf{Z}(\theta))] &= ([I_n] + s^2[\mathbf{Z}(\theta)]^T[M]\,[\mathbf{Z}(\theta)])^{-1/2}\,,\\
[\mathbf{W}(\theta)] &= ([V] + s\,[\mathbf{Z}(\theta)])\,[H_s(\mathbf{Z}(\theta))]\,.
\end{aligned}
$$

3.4. Analysis of some mathematical properties of the stochastic reduced-order basis

From Eqs. (3.19) to (3.24), it follows that the random matrix $[\mathbf{W}]$ is a nonlinear function of $\mathbf{G}$ that can be expected in general to be a non-Gaussian second-order random matrix. Since it depends on the hyperparameter vector $\boldsymbol{\alpha} = (s, \beta, \sigma) \in \mathcal{C}_{\boldsymbol{\alpha}}$, its second-order moments depend on s, β, and $[\sigma]$. The mean value of this SROB is written here as

$$
[\underline{W}(\boldsymbol{\alpha})] = E\{[\mathbf{W}]\}\,.
$$

Its variance, $\mathrm{var}^M_{\mathbf{w}}(\boldsymbol{\alpha})$ (a positive-valued quantity), is defined by

$$
\mathrm{var}^M_{\mathbf{w}}(\boldsymbol{\alpha}) = E\{\|\mathbf{W} - \underline{W}(\boldsymbol{\alpha})\|^2_M\}\,. \tag{3.26}
$$

For all j and j' in $\{1,\ldots,N\}$ and for all k and k' in $\{1,\ldots,n\}$, the components $\mathbb{r}_{jkj'k'}(\boldsymbol{\alpha})$ of the fourth-order correlation tensor $\mathbb{r}(\boldsymbol{\alpha})$ of $[\mathbf{W}]$ are defined by

$$
\mathbb{r}_{jkj'k'}(\boldsymbol{\alpha}) = \frac{E\{[\mathbf{W} - \underline{W}(\boldsymbol{\alpha})]_{jk}[\mathbf{W} - \underline{W}(\boldsymbol{\alpha})]_{j'k'}\}}{\sqrt{E\{[\mathbf{W} - \underline{W}(\boldsymbol{\alpha})]^2_{jk}\}E\{[\mathbf{W} - \underline{W}(\boldsymbol{\alpha})]^2_{j'k'}\}}}\,.
$$

Next, some important properties of the mean value of the SROB $[\mathbf{W}]$, its variance, and its correlation tensor are highlighted.

Let $[\mathbf{Z}]$ be the random matrix defined by Eqs. (3.21)–(3.24). In Appendix A, it is proved that the mean value $[\underline{W}(s,\beta,\sigma)]$ of $[\mathbf{W}]$ (defined by Eqs. (3.19)–(3.24)) is such that

$$
[\underline{W}(s,\beta,\sigma)] = [V]\,E\{[H_s(\mathbf{Z})]\} \quad , \quad [B]^T[\underline{W}(s,\beta,\sigma)] = [0_{N_{\mathrm{CD}},n}]\,. \tag{3.27}
$$

In general, for all $s > 0$, $\beta > 0$, and $[\sigma]$ in $\mathbb{M}^u_n$, $E\{[H_s(\mathbf{Z})]\} \neq [I_n]$. This means that for all $s > 0$, $\beta > 0$, and $[\sigma] \in \mathbb{M}^u_n$,

$$
[\underline{W}(s,\beta,\sigma)] \neq [V] \quad , \quad [\underline{W}(s,\beta,\sigma)]^T[M]\,[\underline{W}(s,\beta,\sigma)] \neq [I_n]\,,
$$

and for $s = 0$,

$$
[\underline{W}(0,\beta,\sigma)] = [V] \quad , \quad [\underline{W}(0,\beta,\sigma)]^T[M]\,[\underline{W}(0,\beta,\sigma)] = [I_n] \quad , \quad \forall\,[\sigma] \in \mathbb{M}^u_n\,.
$$

In Appendix B, it is proved that

$$
\|\underline{W}(s,\beta,\sigma)\|^2_M \leq n \quad , \quad \lim_{s\to+\infty} \|\underline{W}(s,\beta,\sigma)\|^2_M = 0 \quad , \quad \forall\,\beta > 0 \quad , \quad \forall\,[\sigma] \in \mathbb{M}^u_n\,. \tag{3.28}
$$

In the same Appendix B, it is also proved that

$$
\mathrm{var}^M_{\mathbf{w}}(0,\beta,\sigma) = 0 \quad , \quad \lim_{s\to+\infty} \mathrm{var}^M_{\mathbf{w}}(s,\beta,\sigma) = n \quad , \quad \forall\,\beta > 0\ ,\ \forall\,[\sigma] \in \mathbb{M}^u_n\,, \tag{3.29}
$$

and the mapping $s \mapsto \mathrm{var}_{\mathbf{w}}(s,\beta,\sigma)$ is an increasing function that is asymptotic to n. These results show that the statistical fluctuations of the SROB are bounded and cannot grow indefinitely. This is also due to the constraint $[\mathbf{W}]^T[M]\,[\mathbf{W}] = [I_n]$ almost surely.

Furthermore, it can be easily proved that for all $s > 0$, $\beta > 0$, and $[\sigma] \in \mathbb{M}_n^u$, the correlation tensor $\{\mathbb{r}_{jkj'k'}(s,\beta,\sigma)\}_{jkj'k'}$ can be neither written as $a_{jj'}\delta_{kk'}$, neither as $\delta_{jj'}b_{kk'}$, nor as $\delta_{jj'}\delta_{kk'}$. This implies that all components of the tensor $\mathbb{r}(s,\beta,\sigma)$ are not equal to zero — specifically, there are N^2n^2 non zero components.

Finally, it is proved in Appendix C that for all $\beta > 0$ and $[\sigma]$ in $\mathbb{M}_n^u$,

$$\lim_{s\to+\infty} \ll [\mathbf{W}], [V] \gg_M = 0 \quad \text{almost surely.} \tag{3.30}$$

3.5. Continuous interpretation of the stochastic reduced-order basis

A continuous interpretation of the SROB $[\mathbf{W}]$ constructed in Section 3.3 is given in Appendix D. The generator of the random matrix $[\mathbf{G}(\beta)]$ is also described in details in Section D.5 of Appendix D. For all numerical applications presented in this paper, the parameter ν_p introduced in Eq. (D.41) of Appendix D for the generation of the random matrix $[\mathbf{G}(\beta)]$ is set to 20.

4. PERFORMANCE ANALYSIS FOR A SIMPLE EXAMPLE

Here, various performance aspects of the SROB concept and proposed approach for constructing it, including the ability of this SROB to control the statistical fluctuations of the QoI, are assessed through a first, simple, numerical example. Because the explored issues do not necessarily depend on any parametric variation, nonlinear behavior, or dynamic aspect of the modeled system of interest, a generic linear static system is considered for this purpose.

4.1. Generic linear static problem

The considered linear static problem has $N = 1\,000$ dofs. For this problem, Eq. (2.1) is rewritten as

$$[K]\,\mathbf{y} = \mathbf{f}\,, \tag{4.1}$$

where $\mathbf{y} = (y_1, \ldots, y_N)$ is the dimensionless displacement vector belonging to $\mathbb{R}^N$, $[K]$ belongs to $\mathbb{M}_N^+$, and $\mathbf{f} = (f_1, \ldots, f_N)$ is the dimensionless force vector belonging to $\mathbb{R}^N$ and satisfying $f_1 = f_N = 0$. The generation of all of these quantities is described in Appendix E so that the interested reader can reproduce this example. Equation (4.1) has a unique solution that satisfies $y_1 = y_N = 0$ (Dirichlet boundary conditions) and

$$[B]^T\mathbf{y} = \mathbf{0}_{N_{\text{CD}}} \quad , \quad \text{where} \quad [B]^T[B] = [I_N]\,. \tag{4.2}$$

The matrix $[B]$ belongs to $\mathbb{M}_{N,N_{\text{CD}}}$, where $N_{\text{CD}} = 2$. Its construction is also described in Appendix E.

For this simple example, the QoI (see Eq. (2.3)) is defined as

$$\mathbf{o}^D = \mathbf{y}\,,$$

where $\mathbf{y}$ is the unique solution of Eq. (4.1). Hence, $\mathbf{o}^D = (o_1^D, \ldots, o_N^D)$ belongs to $\mathbb{R}^{m_\mathbf{o}}$ with $m_\mathbf{o} = N$, and the superscript "D" stands for "Displacement".

Given that the problem described above is a generic, contrived problem, Appendix E also describes the perturbation approach followed to generate a surrogate set of "experimental" data.

4.2. Reduced-order model

Let $n = 20$, and let $\{\boldsymbol{\varphi}^1, \ldots, \boldsymbol{\varphi}^n\}$ denote the first n eigenvectors associated with the first n eigenvalues $0 < \lambda_1 < \ldots < \lambda_n$ of the matrix $[K]$. Hence,

$$[K]\,\boldsymbol{\varphi}^k = \lambda_k\,\boldsymbol{\varphi}^k \quad , \quad <\boldsymbol{\varphi}^k\,,\boldsymbol{\varphi}^{k'}> = \delta_{kk'}\,. \tag{4.3}$$

Setting $[M] = [I_N]$, the ROB $[V]$ in $\mathbb{M}_{N,n}$ is constructed as $[V] = [\boldsymbol{\varphi}^1 \ldots \boldsymbol{\varphi}^n]$. From Eqs. (4.1) and (4.2) (or the construction of $[K]$ given in Appendix E), it follows that

$$[B]^T[V] = [0_{N_{\text{CD}},n}]\,.$$

For problem (4.1), the projection-based ROM associated with the ROB $[V]$ can be written as (see Eqs. (2.6) and (2.7))

$$\mathbf{y}^{(n)} = [V]\,\mathbf{q}\,, \tag{4.4}$$
$$[V]^T[K]\,[V]\,\mathbf{q} = [V]^T\mathbf{f}\,.$$

The corresponding approximation of the QoI, $\mathbf{o}^{(D,n)} = (o_1^{(D,n)}, \ldots, o_N^{(D,n)})$, is (see Eq. (2.9))

$$\mathbf{o}^{(D,n)} = \mathbf{y}^{(n)}\,. \tag{4.5}$$

At this point, it is noted that n is set here to $n = 20$ so that a modeling error (truncation error) is generated by using the chosen ROM instead of the HDM for solving problem (4.1). Indeed, for $n = 20$, $\mathbf{f}$ does not belong to the subspace spanned by $\{\boldsymbol{\varphi}^1, \ldots, \boldsymbol{\varphi}^n\}$ (see Appendix E).

A second QoI is introduced in the framework of this simple numerical example, namely, the vector $\mathbf{o}^{(E,n)} = (o_1^{(E,n)}, \ldots, o_n^{(E,n)})$ in $\mathbb{R}^n$ defined as

$$\mathbf{o}^{(E,n)} = \boldsymbol{\lambda}^{(n)}\,, \tag{4.6}$$

where $\boldsymbol{\lambda}^{(n)} = (\lambda_1, \ldots, \lambda_n)$ is the vector of the first n eigenvalues $0 < \lambda_1 \leq \ldots \leq \lambda_n$ of $[K]$. To this effect, it is also noted that the ROM approximation of the eigenvalue problem (4.3) can be rewritten, for $k = 1, \ldots, n$, as $\boldsymbol{\varphi}^k = [V]\,\mathbf{q}^k$, where

$$[V]^T[K]\,[V]\,\mathbf{q}^k = \lambda_k\,\mathbf{q}^k\,.$$

For this simple example, the family $\{\mathbf{q}^1, \ldots, \mathbf{q}^n\}$ constitutes the canonical basis of $\mathbb{R}^n$. This observation is introduced here in order to analyze the capability of the nonparametric probabilistic approach proposed in this paper to generate statistical fluctuations for the random eigenvalues.

4.3. Predictions performed using the high-dimensional and reduced-order models

Figure 1 (left) contrasts, for the dimensionless displacements, the graph of $j \mapsto o_j^D$ computed using the HDM and that of $j \mapsto o_j^{(D,n)}$ computed using the ROM. The effect of the truncation error is noticeable. Figure 1 (right) contrasts, for the first n eigenvalues of $[K]$, the graph of $k \mapsto o_k^E$ computed using the HDM and its counterpart computed using the ROM. For this problem, due to the construction of the matrix $[K]$ described in Appendix E and that of the ROB, the eigenvalues λ_k for $k = 1, \ldots, n$ are the same for the HDM and its ROM — that is, $\mathbf{o}^E = \mathbf{o}^{(E,n)}$.

4.4. Stochastic reduced-order model

From Eqs. (2.13) to (2.16), it follows that the SROM associated with Eqs. (4.4) to (4.5) is obtained by substituting $[V]$ with the random matrix $[\mathbf{W}]$. The corresponding quantities $\mathbf{y}^{(n)}$, $\mathbf{q}$, and $\mathbf{o}^{(n)}$ are the random vectors $\mathbf{Y}^{(n)}$, $\mathbf{Q}$, and $\mathbf{O}^{(D,n)}$. Specifically, the SROM is given by

$$\mathbf{Y}^{(n)} = [\mathbf{W}]\,\mathbf{Q}\,,$$
$$[\mathbf{W}]^T[K]\,[\mathbf{W}]\,\mathbf{Q} = [\mathbf{W}]^T\mathbf{f}\,,$$
$$\mathbf{O}^{(D,n)}(\boldsymbol{\alpha}) = \mathbf{Y}^{(n)}\,.$$

For the eigenvalue problem, the vector $\boldsymbol{\lambda}^{(n)}$ becomes the random vector $\boldsymbol{\Lambda}^{(n)} = (\Lambda_1, \ldots \Lambda_n)$, the SROM for the eigenvalue problem is

$$[\mathbf{W}]^T[K]\,[\mathbf{W}]\,\mathbf{Q}^k = \Lambda_k\mathbf{Q}^k\,,$$

and the QoI $\mathbf{o}^{(E,n)}$ defined in Eq. (4.6) becomes the random vector

$$\mathbf{O}^{(E,n)}(\boldsymbol{\alpha}) = \boldsymbol{\Lambda}^{(n)}\,.$$

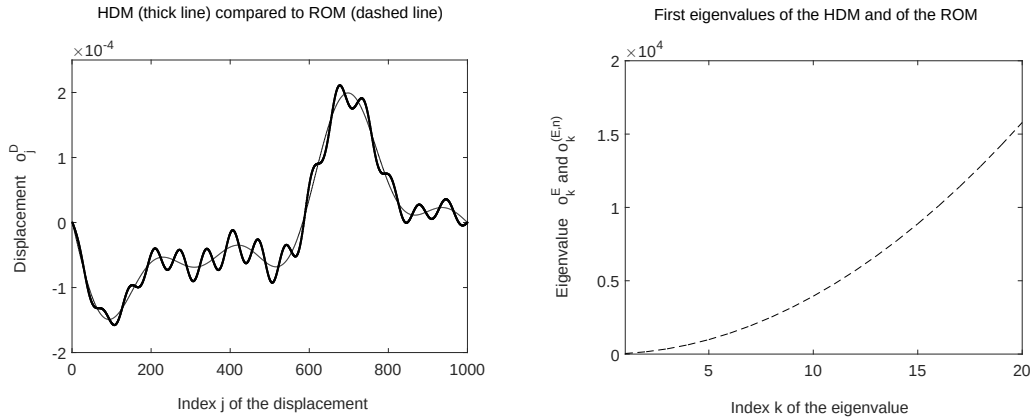


Figure 1. Left figure: Dimensionless displacements, graph of $j \mapsto o_j^D$ computed using the HDM (thick line) and counterpart $j \mapsto o_j^{(D,n)}$ computed using the ROM (dashed line) — Right figure: First n eigenvalues, graph of $k \mapsto o_k^E$ computed using the HDM (dashed line) and counterpart $k \mapsto o_k^{(E,n)}$ computed using the ROM (dashed line).

4.5. Identification of the hyperparameter vector

For this problem, the hyperparameter vector $\boldsymbol{\alpha} = (s, \beta, \sigma)$ defined in Section 3.3.3 has the length $m_{\boldsymbol{\alpha}} = 2 + n(n+1)/2 = 212$. It is identified using the adaptation to $m_\mu = 1$ of the nonlinear LS method presented in Section 2.4. Specifically, $J_{\text{mean}}(\boldsymbol{\alpha})$ and $J_{\text{std}}(\boldsymbol{\alpha})$ are defined in this case as

$$J_{\text{mean}}(\boldsymbol{\alpha}) = \|\mathbf{o}^{\text{ref}} - E\{\mathbf{O}^{(n)}(\boldsymbol{\alpha})\}\|^2 / \|\mathbf{o}^{\text{ref}}\|^2 , \tag{4.7}$$

$$J_{\text{std}}(\boldsymbol{\alpha}) = \|\mathbf{v}^{(\text{ref},n)} - \mathbf{v}^{(n)}(\boldsymbol{\alpha})\|^2 / \|\mathbf{v}^{(\text{ref},n)}\|^2 , \tag{4.8}$$

where $\mathbf{v}^{(\text{ref},n)} = (v_1^{(\text{ref},n)}, \ldots, v_{m_{\mathbf{o}}}^{(\text{ref},n)})$ is such that,

$$v_j^{(\text{ref},n)} = \gamma\, |o_j^{\text{ref}} - o_j^{(n)}| \quad , \quad j = 1, \ldots m_{\mathbf{o}} , \tag{4.9}$$

and $\gamma > 0$ is varied in the computations reported in the following sections. In Eq. (4.8), $\mathbf{v}^{(n)}(\boldsymbol{\alpha}) = (v_1^{(n)}(\boldsymbol{\alpha}), \ldots, v_{m_{\mathbf{o}}}^{(n)}(\boldsymbol{\alpha}))$ is such that

$$v_j^{(n)}(\boldsymbol{\alpha}) = \{\, E\{O_j^{(n)}(\boldsymbol{\alpha})^2\} - (E\{O_j^{(n)}(\boldsymbol{\alpha})\})^2\} \,\}^{1/2} \quad , \quad j = 1, \ldots m_{\mathbf{o}} .$$

The reader is reminded that in Eqs. (4.7) and (4.8), $\mathbf{o}^{\text{ref}}$ is the target for the mean value with a weight w_J and $\mathbf{v}^{(\text{ref},n)}$ is the target for the standard deviation with a weight $1 - w_J$ (see Eq. (2.17)).

Throughout the remainder of the discussion of the academic problem examined here, the following cases will be considered for the vector-valued QoI $\mathbf{o}^{\text{ref}}$:

$\diamond$ $\mathbf{o}^{\text{ref}} = \mathbf{o}^D = \mathbf{y}$, with $m_{\mathbf{o}} = N$, which is the deterministic displacement vector computed using the HDM and does not coincide with the deterministic approximation $\mathbf{y}^{(n)}$ computed using the ROM. The random vector $\mathbf{O}^{(n)}(\boldsymbol{\alpha})$ is then chosen as $\mathbf{O}^{(D,n)}(\boldsymbol{\alpha})$. For this choice, the constructed stochastic model of uncertainties takes into account the truncation error induced by the ROM.

$\diamond$ $\mathbf{o}^{\text{ref}} = \mathbf{o}^{D,\text{exp}} = \mathbf{y}^{\text{exp}}$, with $m_{\mathbf{o}} = N$, which is the experimentally measured (surrogate) displacement vector (see Appendix E) and does not coincide with the prediction $\mathbf{y}$ performed using the HDM. For this choice, the constructed stochastic model of uncertainties takes into account both the truncation error induced by the ROM and the modeling errors (with respect to the surrogate real system) introduced during the construction of the underlying HDM. The random vector $\mathbf{O}^{(n)}(\boldsymbol{\alpha})$ is chosen as $\mathbf{O}^{(D,n)}(\boldsymbol{\alpha})$.

$\diamond$ $\mathbf{o}^{\text{ref}} = \mathbf{o}^{E,\text{exp}} = \boldsymbol{\lambda}^{\text{exp}}$, with $m_{\mathbf{o}} = n$, which is the experimentally measured set of the first n eigenvalues and does not coincide with the deterministic counterpart $\boldsymbol{\lambda}^{(n)}$ computed using the

ROM (or in this case, the HDM), and the random vector $\mathbf{O}^{(n)}(\boldsymbol{\alpha})$ is chosen as $\mathbf{O}^{(E,n)}(\boldsymbol{\alpha})$. This choice allows the evaluation of the potential of the proposed nonparametric probabilistic approach for simulating different types of errors induced by the ROM and/or underlying HDM.

It is noted here that the optimization problem

$$\boldsymbol{\alpha}^{\text{opt}} = (s^{\text{opt}}, \beta^{\text{opt}}, [\sigma^{\text{opt}}]) = \min_{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\alpha}}} J(\boldsymbol{\alpha}) \tag{4.10}$$

is not a convex problem. Therefore, only an approximation of $\boldsymbol{\alpha}^{\text{opt}}$ can be computed in general (recall that $m_{\boldsymbol{\alpha}} = 212$). Many numerical tests were performed to analyze the roles played by the components s, β, and $[\sigma]$ of $\boldsymbol{\alpha}$. These tests have led to shaping the final optimization algorithm presented in Appendix F. Nevertheless, this algorithm can be replaced by any other preferred solution algorithm for problem (4.10).

4.6. Sensitivity analysis of the stochastic model with respect to model reduction errors

For the purpose of performing a numerical sensitivity analysis of the stochastic model with respect to the modeling error induced by model reduction, the QoI is set here to $\mathbf{o}^{\text{ref}} = \mathbf{o}^D = \mathbf{y}$ with $m_{\mathbf{o}} = N$. The weight w_J is set to $w_J = 0.9$, so that the target $\mathbf{o}^{\text{ref}}$ for the mean value is preponderant for the standard deviation (see Eq. (4.9)). The optimization problem (4.10) is solved using the interior-point algorithm described in Appendix F with $\varepsilon_0 = 0.01$, $\beta_d = 0.01$, $\beta_u = 0.3$, and $\sigma_u = 20$, and the initialization $s_0 = 0.05$, $\beta_0 = 0.2$, and $[\sigma_0] = [I_n]$. Furthermore, the solution $\boldsymbol{\alpha}^{\text{opt}} = (s^{\text{opt}}, \beta^{\text{opt}}, [\sigma^{\text{opt}}])$ is computed for $\gamma = 0.2, 0.5, 0.8$, and 1.0. The Monte Carlo method is used as the stochastic solver with $1\,000$ independent realizations, and the mean-square convergence is reached with a reasonable accuracy.

Figure 2 (left) displays the graphs of four arbitrary trajectories (realizations) of the discrete random field $j \mapsto O_j^{(D,n)}$ computed using the SROM for $\gamma = 0.8$. The reader can observe that the Dirichlet boundary conditions are preserved, and that the computed trajectories are as regular as predicted by the theory (see Appendix D).

Figure 2 (right) reports the graph of the function $k \mapsto [c_n^{\text{opt}}]_{kk}$ (diagonal entries), where $[c_n^{\text{opt}}] = [\sigma^{\text{opt}}]^T[\sigma^{\text{opt}}]$ corresponds to the optimal value $[\sigma^{\text{opt}}]$ of $[\sigma]$ (see Eq. (3.16)). For $\gamma = 0.2, 0.5, 0.8$, and 1.0, the optimal values of s and β are $s^{\text{opt}} = 0.0111, 0.0158, 0.0120$, and 0.0173, and $\beta^{\text{opt}} = 0.071$, 0.036, 0.207, and 0.111.

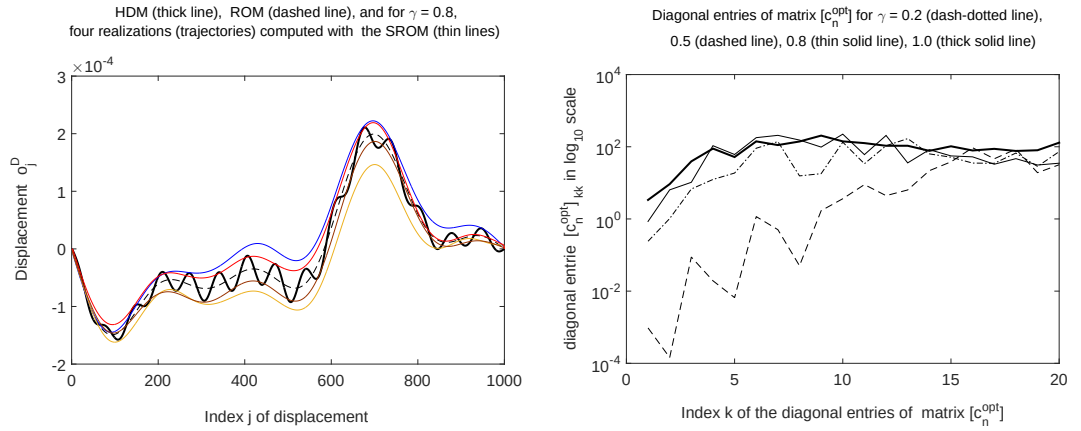


Figure 2. Left figure: Four sample realizations of the discrete random field $j \mapsto O_j^{(D,n)}$ for $\gamma = 0.8$ (four colored thin lines) — Right figure: Graph of $k \mapsto [c^{\text{opt}}]_{kk}$ (in $\log_{10}$ scale) for $\gamma = 0.2$ (dash-dotted line), 0.5 (dashed line), 0.8 (thin solid line), and 1.0 (thick solid line).

Figure 3 (left) displays the graph of $j \mapsto o_j^D$ corresponding to the HDM target, that of $j \mapsto o_j^{(D,n)}$ computed using the ROM and $\gamma = 0.2, 0.5, 0.8$, and 1.0, and the mean value of the discrete random

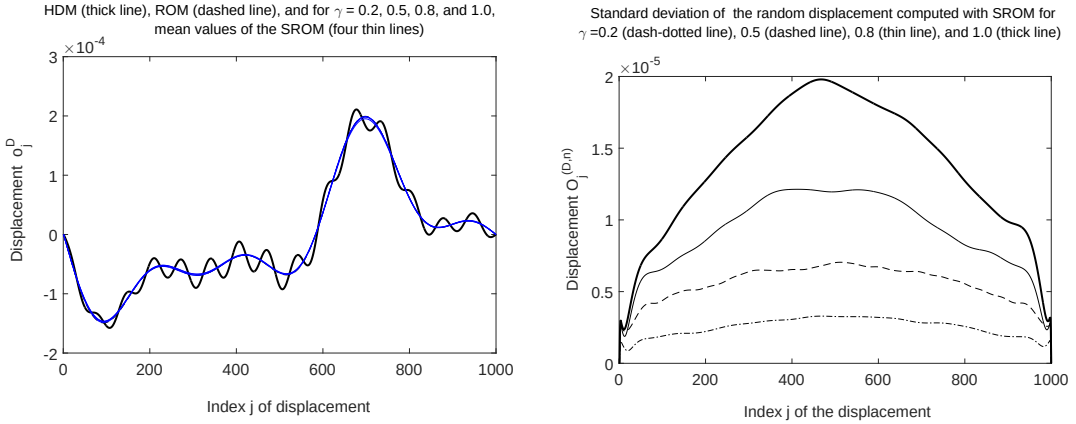


Figure 3. Left figure: Graph of $j \mapsto o_j^D$ of the HDM target (black line), graph of $j \mapsto o_j^{(D,n)}$ computed using the ROM for $\gamma = 0.2$, 0.5, 0.8, and 1.0, and mean value of the random field $j \mapsto O_j^{(D,n)}$ computed using the SROM (blue thick lines superimposed on the black line) — Right figure: Standard deviation of $j \mapsto O_j^{(D,n)}$ computed using the SROM for $\gamma = 0.2$ (dashed-dotted line), 0.5 (dashed line), 0.8 (thin solid line), and 1.0 (thick solid line).

field $j \mapsto O_j^{(D,n)}$ constructed with the SROM. The ROM and mean value of the SROM are almost coinciding for all considered values of γ. Figure refFIG3 (right) displays, for the different considered values of γ, the standard deviation of $j \mapsto O_j^{(D,n)}$ computed using the SROM. As it can be expected, the standard deviation increases with the amplitude γ of its target value.

In Figure 4, the four displayed figures summarize the results obtained for the numerical sensitivity analysis of the stochastic model with respect to the error induced by using the ROM instead of the HDM for solving problem (4.1). Each figure pertains to a different value of γ (0.2, 0.5, 0.8, and 1.0) and displays, for the dimensionless displacements, the graph of $j \mapsto o_j^D$ corresponding to the HDM target, the graph of $j \mapsto o_j^{(D,n)}$ computed using the ROM, and the confidence region with a probability $p_c = 0.98$ of the discrete random field $j \mapsto O_j^{(D,n)}$ constructed using the SROM. The upper envelope of the confidence region corresponds to the quantile for the probability p_c, and the lower envelope to the quantile for the probability $1 - p_c$. These figures show that the statistical fluctuations increase as the amplitude γ of the target for the standard deviation increases (as expected). They also show that the confidence region is relatively well-centered around the response computed using the ROM, which is close to the average value of the HDM target (see Figure 3) and was chosen as the target for the mean value. Finally, the reader can observe that the prediction performed for $\gamma = 0.8$ is simply excellent.

4.7. Performance in the presence of modeling errors in both of the high-dimensional and reduced-order models

For the purpose of assessing the ability of the proposed nonparametric stochastic model to account for modeling errors introduced at both HDM and ROM levels, the QoI is set to the displacement $\mathbf{o}^D = \mathbf{y}$ with $m_{\mathbf{o}} = N$, the target of the mean value is set to $\mathbf{o}^{\text{ref}} = \mathbf{o}^{D,\text{exp}} = \mathbf{y}^{\text{exp}}$ — which corresponds to the surrogate experimental data and is denoted as "EXP target" in the figures to follow. The graph of $j \mapsto o_j^{D,\text{exp}}$ is displayed in Figure 6 (top left). The target $\mathbf{v}^{(\text{ref},n)}$ of the standard deviation is defined in Eq. (4.9), where the amplitude is controlled by the parameter γ. (It is noted here that the task of generating an SROM for which the mean value of the stochastic response will be close to the average of the EXP target and the confidence region of the statistical fluctuations will contain this EXP target is a difficult one). The weight $w_J = 0.9$ is chosen so that the target $\mathbf{o}^{\text{ref}}$ of the mean value is preponderant during the identification of the hyperparameter $\boldsymbol{\alpha}$. This identification is performed in

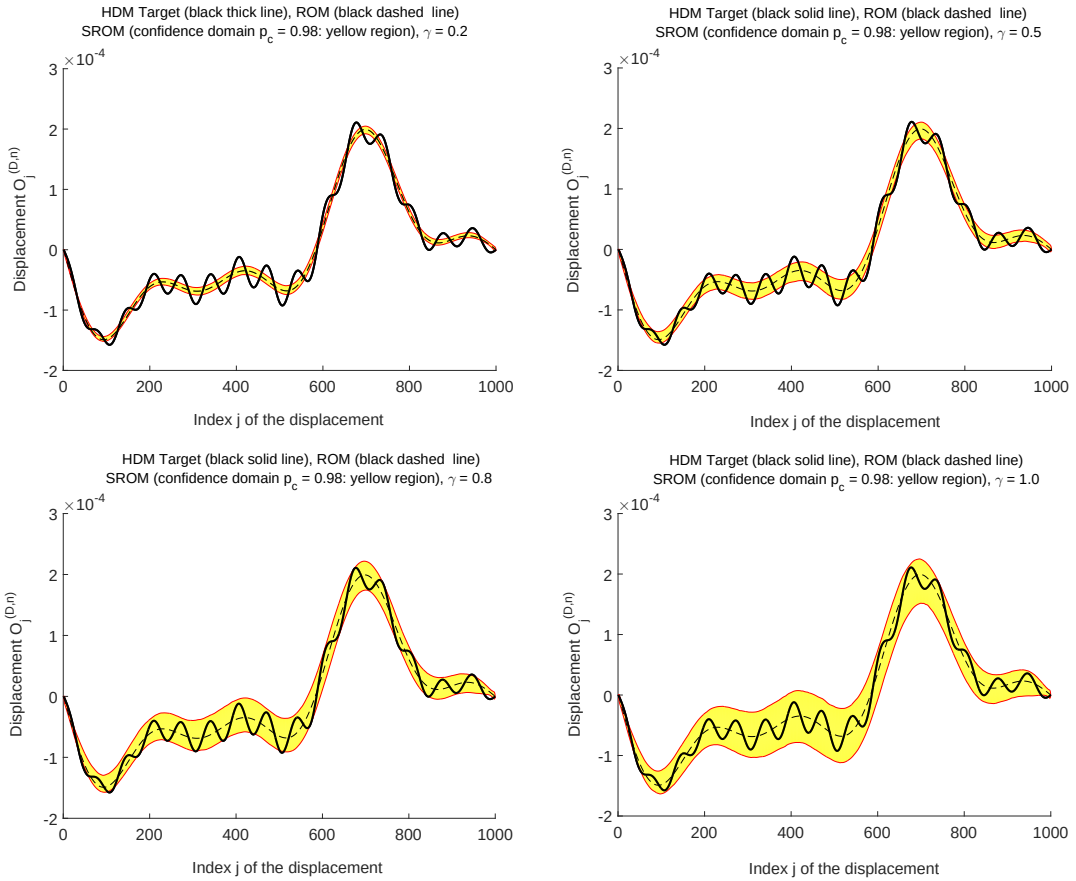


Figure 4. For the dimensionless displacements, graph of $j \mapsto o_j^D$ of the HDM target (black thick solid line), graph of $j \mapsto o_j^{(D,n)}$ computed using the ROM (black dashed line), and confidence region for $p_c = 0.98$ (yellow region with red upper and lower envelopes) computed for $j \mapsto O_j^{(D,n)}$ using the SROM and $\gamma = 0.2$ (top left), $\gamma = 0.5$ (top right), $\gamma = 0.8$ (bottom left), and $\gamma = 1.0$ (bottom right).

this case for 3 different values of γ, namely, $\gamma = 0.1$, 0.2, and 0.3. The optimization problem (4.10) is solved using the interior-point algorithm described in Appendix F with $\varepsilon_0 = 0.01$, $\beta_d = 0.0005$, $\beta_u = 0.01$, and $\sigma_u = 20$, and the initialization $s_0 = 0.011$, $\beta_0 = 0.001$, and $[\sigma_0] = [I_n]$. As in the previous analysis, the Monte Carlo method is used as the stochastic solver with $1\,000$ independent realizations (and the mean-square convergence is reached with a reasonable accuracy).

Figure 5 (left) displays the graphs of four arbitrary trajectories (realizations) of the discrete random field $j \mapsto O_j^{(D,n)}$ computed using the SROM and $\gamma = 0.3$. As in Section 4.6, the Dirichlet boundary conditions are preserved and the trajectories are regular.

Figure 5 (right) displays the graph of the function $k \mapsto [c_n^{\text{opt}}]_{kk}$ (diagonal entries), where $[c_n^{\text{opt}}] = [\sigma^{\text{opt}}]^T[\sigma^{\text{opt}}]$ corresponds to the optimal value $[\sigma^{\text{opt}}]$ of $[\sigma]$ (see Eq. (3.16)). For $\gamma = 0.1$, 0.2, and 0.3, the optimal values of s and β are $s^{\text{opt}} = 0.0138$, 0.0142, and 0.0275, and $\beta^{\text{opt}} = 0.0005$, 0.0074, and 0.0077.

Figure 6 (top left) displays the graphs of $j \mapsto o_j^{D,\text{exp}}$ of the EXP target, $j \mapsto o_j^D$ computed using the HDM, $j \mapsto o_j^{(D,n)}$ computed using the ROM and $\gamma = 0.1$, 0.2, and 0.3, and the mean value of the discrete random field $j \mapsto O_j^{(D,n)}$ constructed using the SROM. The reader can observe that the solution computed using the HDM is different from the EXP target, which simulates model uncertainties in the HDM, and that the ROM yields an averaging of the solution predicted by the HDM. The reader can also observe that the mean values are close to the EXP target, which is

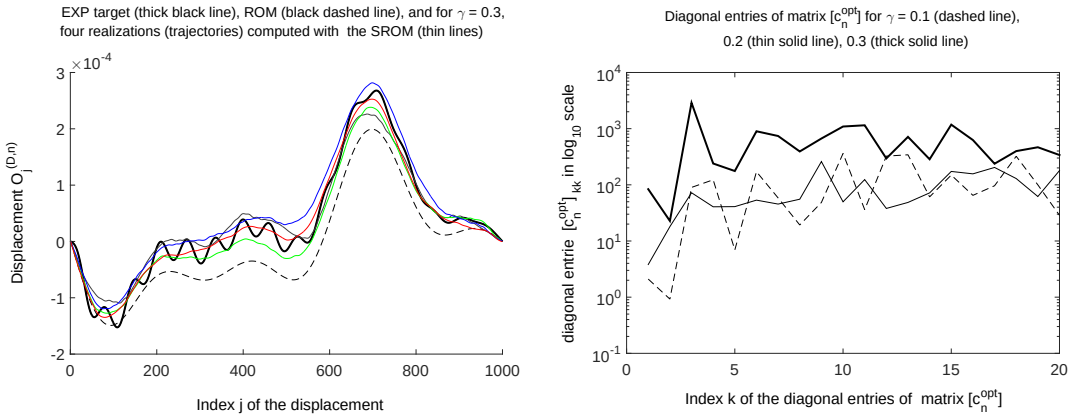


Figure 5. Left figure: For $\gamma = 0.3$, sample realizations of the discrete random field $j \mapsto O_j^{(D,n)}$ (four colored thin lines) — Right figure: Graph of $k \mapsto [c^{\text{opt}}]_{kk}$ (in $\log_{10}$ scale) for $\gamma = 0.1$ (dashed line), 0.2 (thin line), and 0.3 (thick line).

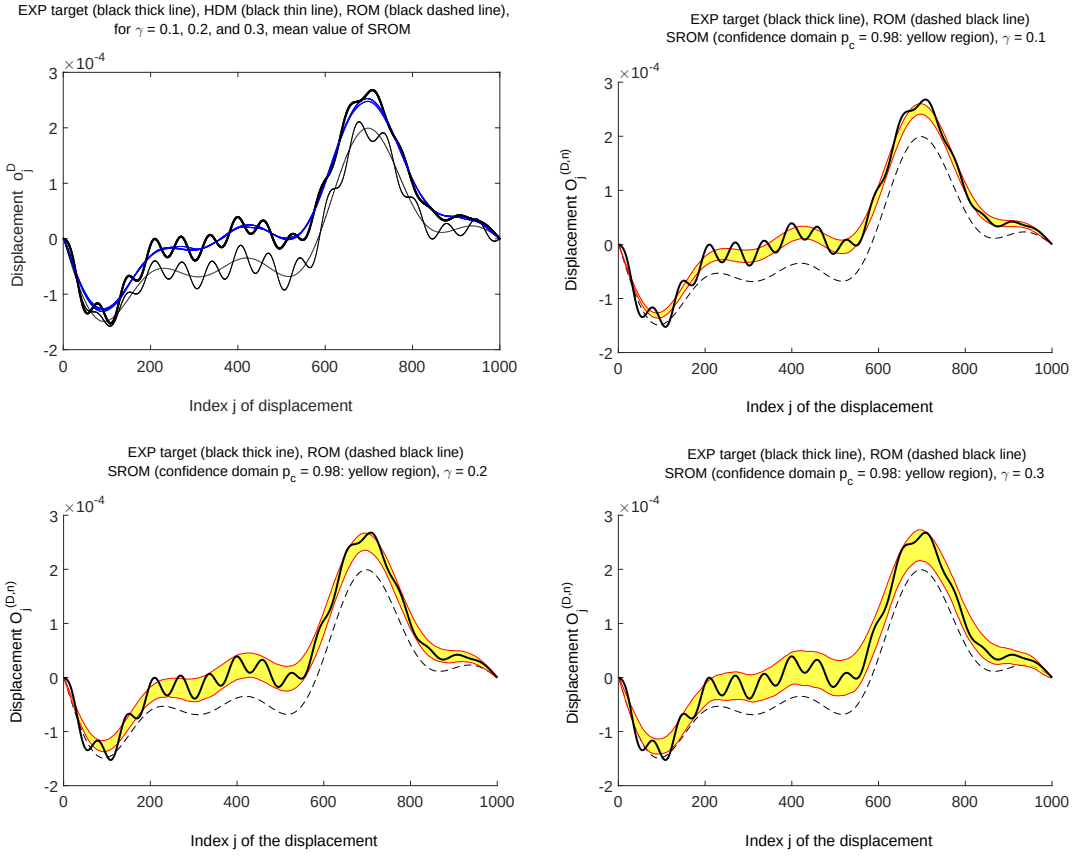


Figure 6. Graphs of $j \mapsto \mathbf{o}_j^{D,\text{exp}}$ (thick solid lines), $j \mapsto \mathbf{o}_j^D$ computed using the HDM (dash-dotted lines), and $j \mapsto o_j^{(D,n)}$ computed using the ROM (dashed lines) — For the discrete random field $j \mapsto O_j^{(D,n)}$ computed using the SROM and $\gamma = 0.1$, 0.2, and 0.3: mean value (three superimposed blue thin lines, top left) — For the confidence region for $p_c = 0.98$ (yellow region with red upper and lower envelopes): $\gamma = 0.1$ (top right), $\gamma = 0.2$ (bottom left), $\gamma = 0.3$ (bottom right).

coherent with the fact that the EXP target is given as the target for the mean value with a weight $w_J = 0.9$. For $\gamma = 0.1$, 0.2, and 0.3, the 3 other figures shown in Figure 6 (top right, bottom left, and bottom right) display the graph of $j \mapsto o_j^{D,\exp}$ of the EXP target, that of $j \mapsto o_j^{(D,n)}$ computed using the ROM, and the confidence region (with a probability $p_c = 0.98$) of the discrete random field $j \mapsto O_j^{(D,n)}$ constructed using the SROM. These figures show that, as expected, the statistical fluctuations increase as the amplitude γ of the target for the standard deviation increases, the domain defined by the confidence region contains the EXP target, and the prediction performed for $\gamma = 0.2$ is excellent.

4.8. Random eigenvalue analysis

Next, the ability of the proposed nonparametric stochastic model to control the mean values and statistical fluctuations of the vector $\mathbf{\Lambda}^{(n)}$ of the first n random eigenvalues $(\Lambda_1, \ldots, \Lambda_n)$ (see (4.4)) associated with the QoI $\mathbf{o}^{(E,n)} = \boldsymbol{\lambda}^{(n)}$ with $m_{\mathbf{o}} = n$ (see (4.6)) is assessed. To this effect, the target of the mean value is set to $\mathbf{o}^{\text{ref}} = \mathbf{o}^{E,\,\exp} = \boldsymbol{\lambda}^{\exp}$, which corresponds to the surrogate experimental data. The target $\mathbf{v}^{(\text{ref},n)}$ of the standard deviation is defined in Eq. (4.9); its amplitude is controlled by the parameter γ. The weight $w_J = 0.9996$ is chosen so that the target $\mathbf{o}^{\text{ref}}$ of the mean value is preponderant during the identification of the hyperparameter $\boldsymbol{\alpha}$. This identification is performed here for $\gamma = 0.14$ and $\gamma = 0.043$. These two values of γ are arbitrarily chosen except for the interesting fact that for one value ($\gamma = 0.14$), the error happens to increase with the rank of the eigenvalues while for the other ($\gamma = 0.043$), the error happens to decrease with the rank of the eigenvalues. The optimization problem is solved using only Stages 1 and 4 of the interior-point algorithm described in Appendix F. For $\gamma = 0.14$, the parameters of this algorithm are set to $\varepsilon_0 = 0.01$, $\beta_d = 0.0001$, $\beta_u = 0.2$, and $\sigma_u = 20$, and the initialization is set to $s_0 = 0.011$, $\beta_0 = 0.01$, and $[\sigma_0]_{kk'} = \delta_{kk'}\, 14 * k/20 + 1$. For $\gamma = 0.043$, they are set to $\varepsilon_0 = 0.01$, $\sigma_u = 20$, $\beta_d = 0.001$ and $\beta_u = 0.003$, and $s_0 = 0.011$, $\beta_0 = 0.002$, and $[\sigma_0]_{kk'} = \delta_{kk'}\, - 14 * k/20 + 15$.

Two targets cases are considered:

$\diamond$ The mean-value target (EXP-target$^+$) and standard-deviation target ($10 \times$ Std-target$^+$) generated for $\gamma = 0.14$, displayed in Figure 7 (left), and corresponding to an increasing of the statistical fluctuations of the random eigenvalues with their index. This choice of targets simulates a modeling error in the HDM that increases the statistical fluctuations of the random eigenvalues with their index (smaller statistical fluctuations for Λ_1 and larger statistical fluctuations for Λ_n).

$\diamond$ The mean-value target (EXP-target$^-$) and standard-deviation target ($10 \times$ Std-target$^-$) generated for $\gamma = 0.043$, displayed in Figure 7 (right), and corresponding to a decreasing of the statistical fluctuations of the random eigenvalues with their index. This choice of targets simulates a modeling error in the HDM that decreases the statistical fluctuations of the random eigenvalues with their index (larger statistical fluctuations for Λ_1 and smaller statistical fluctuations for Λ_n).

In Figure 7, one observes that the two mean-value targets are significatively different from the eigenvalues computed using the HDM (which, as explained earlier, are identical for this specific problem to their counterparts computed using the ROM).

Figure 8 (left) displays the graph of the function $k \mapsto [c_n^{\text{opt}}]_{kk}$ (diagonal entries) where $[c_n^{\text{opt}}] = [\sigma^{\text{opt}}]^T[\sigma^{\text{opt}}]$ corresponds to the optimal value $[\sigma^{\text{opt}}]$ of $[\sigma]$ (see Eq. (3.16)). For $\gamma = 0.14$ and 0.043, the optimal values of s and β are $s^{\text{opt}} = 0.0213$ and 0.0109, and $\beta^{\text{opt}} = 0.0242$ and 0.00298.

For $\gamma = 0.14$ (0.043), which corresponds to an increasing (decreasing) of the statistical fluctuations of the random eigenvalues with their index , the two figures in Figure 9 display the graph of $k \mapsto o_k^{E,\,\exp}$ of the mean-value target, that of $k \mapsto o_k^E = o_k^{(E,n)}$ computed using the HDM, and the confidence region of the random eigenvalues $k \mapsto O_k^{(E,n)}$ constructed by using the SROM for the probability level of $p_c = 0.98$. The reader can observe that whereas the eigenvalues predicted by the HDM are different from the target values, the mean values of the random eigenvalues are close to these target values.

For $\gamma = 0.14$, the statistical fluctuations of the random eigenvalues increase with their index while for $\gamma = 0.043$, they decrease with their index. This is coherent with the imposed standard-deviation targets.

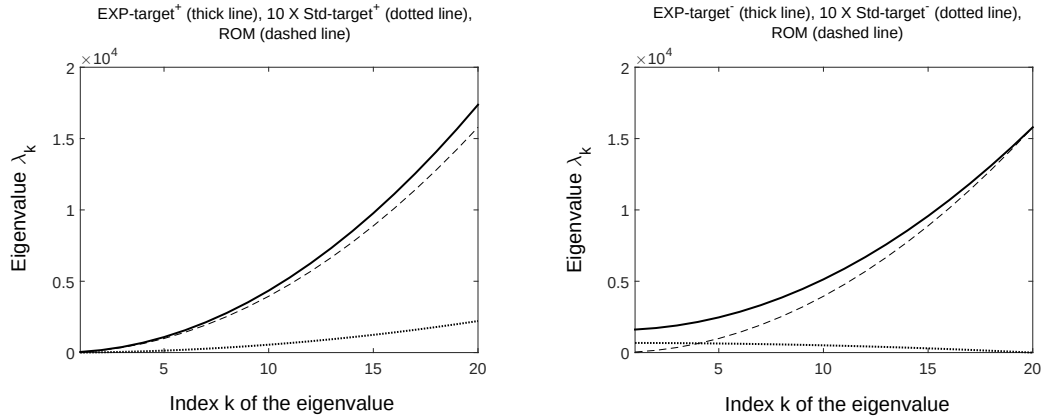


Figure 7. Graphs of $k \mapsto o_k^E = o_k^{(E,n)}$ computed using the HDM and the ROM (dashed lines), $k \mapsto o_k^{E,\,\mathrm{exp}}$ of the mean-value target (thick lines) for $\gamma = 0.14$ (left) and $\gamma = 0.043$ (right), and $k \mapsto v_k^{(\mathrm{ref},n)}$ of the standard-deviation target (dotted lines) for $\gamma = 0.14$ (left) and $\gamma = 0.043$ (right).

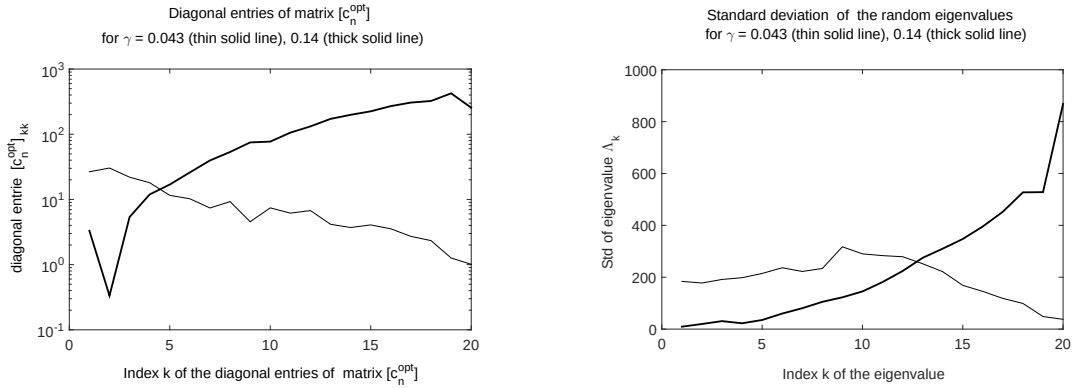


Figure 8. Left figure: Graph of $k \mapsto [c^{\mathrm{opt}}]_{kk}$ for $\gamma = 0.043$ (thin line) and $\gamma = 0.14$ (thick line) — Right figure: Standard deviation of $k \mapsto O_k^{(E,n)}$ computed using the SROM for $\gamma = 0.043$ (thin line) and $\gamma = 0.14$ (thick line).

All reported results demonstrate the capability of the proposed nonparametric probabilistic approach for quantifying model uncertainties to modify the statistical properties of the spectrum of a linear operator and reproduce (in this case surrogate) experimental data.

5. APPLICATION TO A NONLINEAR COMPUTATIONAL STRUCTURAL DYNAMICS PROBLEM

Here, a three-dimensional (3D) dynamic problem associated with a slender, damped, linearly elastic structure with *nonlinear barriers* is considered. The setup of this problem is designed to be sufficiently simple to enable the reproduction of the associated results by the interested reader. On the other hand, constructing an efficient ROM for this application is a difficult task. In the frequency domain, the energy of the considered excitation is concentrated in a narrow frequency band, within an otherwise broader frequency analysis interval. For this reason, the discrepancies between the HDM and ROM predictions are very small in the frequency band of excitation where the uncertainties are small, but larger outside this frequency band where, because of the energy transfer outside the frequency band of excitation due to nonlinearities, the model uncertainties are larger.

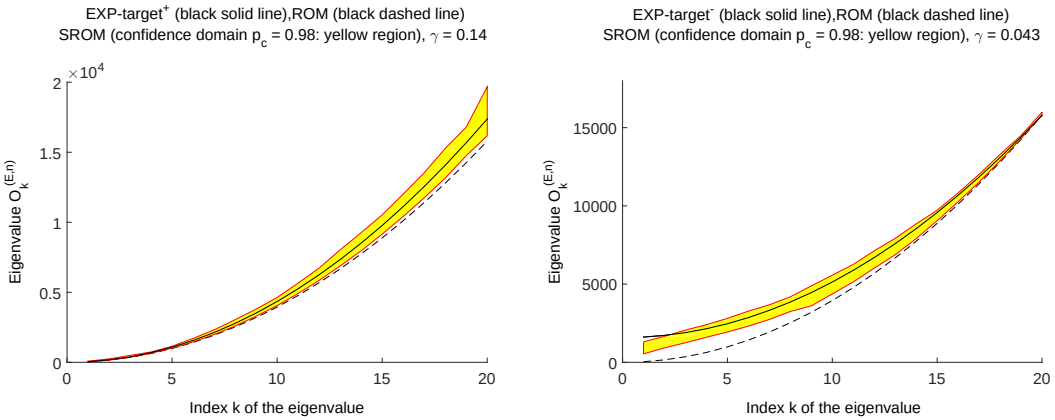


Figure 9. Graphs of $k \mapsto o_k^{E,\,\exp}$ of the mean-value target (thick solid line) and of $k \mapsto \mathbf{o}_k^E = o_k^{(E,n)}$ computed using the ROM (dashed line) — For $p_c = 0.98$, confidence region (yellow region with red upper and lower envelopes) of the random eigenvalues $k \mapsto O_k^{(E,n)}$ computed using the SROM: $\gamma = 0.14$ (left), $\gamma = 0.043$ (right).

Consequently, such uncertainties are associated with second-order contributions to the solution, and the main purpose of this problem is to demonstrate the ability of the proposed nonparametric probabilistic approach to predict them.

5.1. Description of a mechanical system

The mechanical system considered here is a 3D linear elastic structure with two elastic barriers that induce impact nonlinearities. It is defined in a cartesian coordinate system $Ox_1x_2x_3$ (see Figure 10). Its cylindrical geometry has a length $L_1 = 1.2\,m$, and a rectangular section with height $L_2 = 0.12\,m$ and width $L_3 = 0.24\,m$. The two end sections are located at $x_1 = 0$ and $x_1 = L_1$. The origin O is in the corner of the first end section, and Ox_1 is parallel to the main axis of the cylinder. The axis Ox_2 is the transversal axis along the side of length L_2, and Ox_3 is the axis along the side of length L_3 (see Figure 10). The elastic medium is made of a homogeneous and isotropic elastic material for which the Young modulus is $E = 10^{10}\ N/m^2$, the Poisson ratio is $\nu = 0.15$, and the mass density is $\rho = 1500\ Kg/m^3$. Damping is represented using the global damping rate of $\xi_d = 0.01$ for each elastic mode of this structure in the absence of elastic barriers, and is introduced directly at the ROM level. As for the boundary conditions, all displacements are constrained ("locked") along the two lines defined by $\{(x_1, x_2, x_3) : x_1 = 0\,; x_2 = L_2/2\,;\, 0 \le x_3 \le L_3\}$ and by $\{(x_1, x_2, x_3) : x_1 = L_1\,; x_2 = L_2/2\,;\, 0 \le x_3 \le L_3\}$ (see Figure 10).

The elastic barriers induce two nonlinear point forces on the structure: one in the direction of x_2 and applied at the point $(x_1 = 0.66, x_2 = 0, x_3 = L_3/2)$, and another in the direction of x_3 and applied to the point $(x_1 = 0.66, x_2 = L_2/2, x_3 = L_3)$ (see Figures 10 and 12), of intensities $-f_{\text{NL},2}$ and $-f_{\text{NL},3}$, respectively. These intensities, which are assumed to be independent of the velocity, are given by

$$f_{\text{NL},2}(\eta) = k_{b,2}\,(\eta + \varepsilon_{b,2})\,\mathbb{1}_{\mathbb{R}^+}(-\eta - \varepsilon_{b,2}) \quad , \quad \eta \in \mathbb{R}\,, \tag{5.1}$$

$$f_{\text{NL},3}(\eta) = k_{b,3}\,(\zeta - \varepsilon_{b,3})\,\mathbb{1}_{\mathbb{R}^+}(\zeta - \varepsilon_{b,3}) \quad , \quad \zeta \in \mathbb{R}\,, \tag{5.2}$$

where $k_{b,2} = k_{b,3} = 2 \times 10^9\ N/m$ is the elasticity constant of the barriers and $\varepsilon_{b,2} = \varepsilon_{b,3} = 2 \times 10^{-4}\ m$ are the two positive valued gaps.

A time-dependent point force is also applied at the boundary of this slender structure, specifically, at $(x_1 = 0.46, x_2 = 0, x_3 = 0.2)$ (see Figures 10, 12). The x_1 component of this force is zero, and its x_2- and x_3-components are equal to the square integrable real-valued function f_e defined by

$$f_e(t) = \frac{f_0}{\pi\, t}\,\{\sin(t(\omega_c + \Delta\omega_c/2)) - \sin(t(\omega_c - \Delta\omega_c/2)\} \quad , \quad t \in [t_0, T]\,, \tag{5.3}$$

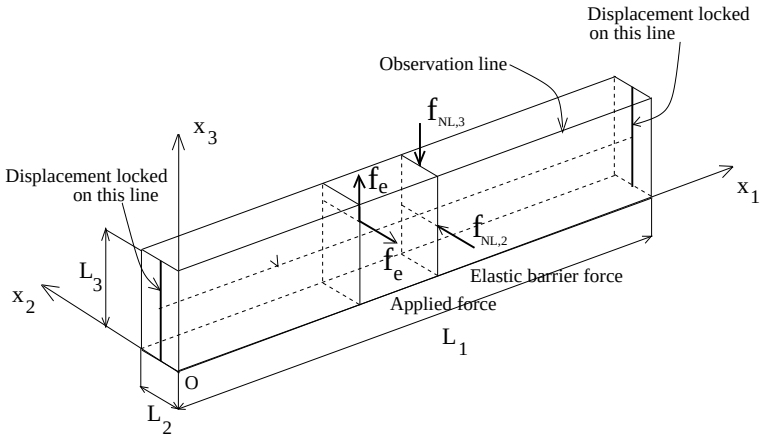


Figure 10. Slender elastic structure with nonlinear barriers.

where $f_0 = 100\ N$, $\omega_c = 2\,\pi \times 470\ rad/s$ is the central circular frequency, and $\Delta\omega_c = 2\,\pi \times 300\ rad/s$ is the circular frequency bandwidth. The graph of $\omega \mapsto |\widehat{f_e}(\omega)|$ defined on $2\,\pi \times [0\,,1550]\ rad/s$, where $\widehat{f_e}(\omega) = \displaystyle\int_{t_0}^{T} e^{-i\omega t} f_e(t)\,dt$, is plotted in Figure 11. The signal energy of the excitation is mainly concentrated in the frequency band $[320\,,620]\ Hz$.

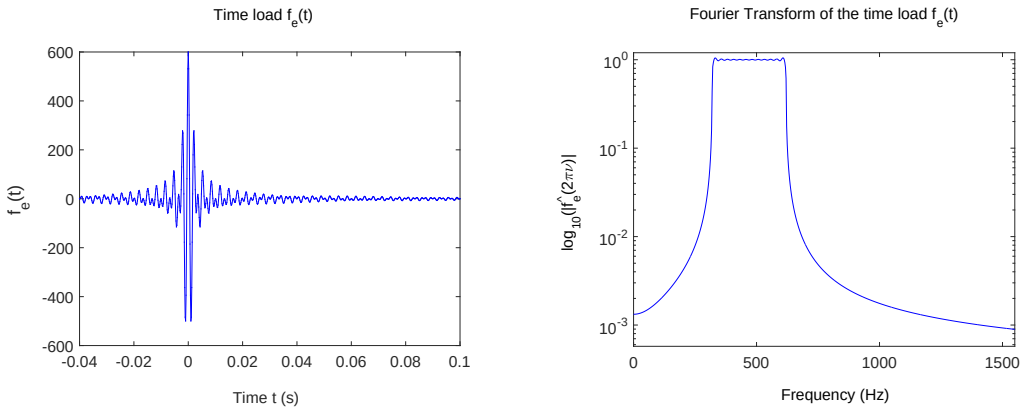


Figure 11. Time-variation of the external point force $t \mapsto f_e(t)$ defined on $[-0.0403\,,0.3790]\ s$ and represented on $[-0.04\,,0.1]\ s$ (left) — Graph of $\omega \mapsto |\widehat{f_e}(\omega)|$ in $\log_{10}$-scale defined on the frequency band of observation $[0\,,1550]\ Hz$ (right).

The line defined by $\{0 \le x_1 \le L_1\,;\, x_2 = 0\,;\, x_3 = L_3\}$ is adopted as an observation line for $t \in [t_0, T]$. In order to limit the number of figures, only the x_2- and x_3-displacements of 2 observation points on this line are considered and denoted here by Obs_{34} and Obs_{51}. Their x_1-coordinates are 0.66 and 1.00, respectively (see their locations in Figure 12). Consequently, only 4 displacement dofs are observed for the purpose of reporting numerical results. The initial time $t_0 < 0$ is written as $t_0 = -m_0\,\pi/\omega_e = -0.0403\ s$, where m_0 is a positive integer set to $m_0 = 50$. At t_0, the system is assumed to be at rest (displacement and velocity fields are zero). The final time $T > 0$ is a positive integer chosen so that the system is returned to the zero equilibrium with a relative error of $\varepsilon_d = 0.1$. Its value is estimated using the equation $\exp(-\xi_d\,\omega_1\,T) = \varepsilon_d$, where $\omega_1 = 2\,\pi \times 96.69\,rad/s$ is the fundamental eigenfrequency of the associated linear undamped dynamical system, which yields $T = 0.3790\ s$. The *frequency band of observation* is chosen as $\mathcal{B}_o = [0\,,\omega_o]$, with $\omega_o = 2\,\pi \times 1550\ rad/s$. The chosen time-interval is sampled at the frequency

$\omega_{\max} = 2\,\pi \times 12,400\ rad/s$. This leads to $10,400$ time-instances and the time-step $\Delta t = \pi/\omega_{\max} = 4.032 \times 10^{-5}\ s$. Similarly, there are $10,400$ frequency steps in the frequency band $[-\omega_{\max}, \omega_{\max}]$ and the sampling frequency step is $\Delta\omega = 2\,\pi \times 2.38\ rad/s$.

5.2. Predictions performed using the high-dimensional model

A 3D finite element model is constructed for the problem described above using a $60 \times 6 \times 12 = 4\,320$ 8-noded solid elements. Hence, this model contains $5\,551$ nodes and $N = 16\,653$ dofs. $N_{\text{CD}} = 78$ Dirichlet boundary conditions are applied at 2×13 nodes (see Figure 12). For this HDM,

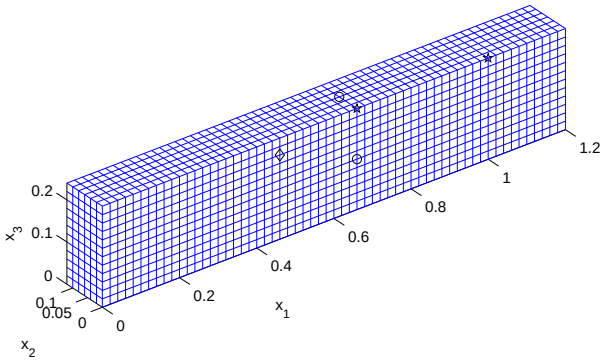


Figure 12. Finite element model: applied force (diamond), two elastic barriers (circles), observation points Obs_{34} and Obs_{51}(pentagrams).

the governing equations (2.1) to (2.3) are rewritten as follows

$$[M]\,\ddot{\mathbf{y}}(t) + \mathbf{g}(\mathbf{y}(t), \dot{\mathbf{y}}(t)) = \mathbf{f}(t) \quad , \quad t \in]t_0, T]\,, \tag{5.4}$$

$$\mathbf{g}(\mathbf{y}(t), \dot{\mathbf{y}}(t)) = [\mathbb{D}]\,\dot{\mathbf{y}}(t) + [K]\,\mathbf{y}(t) + \mathbf{f}_{\text{NL}}(\mathbf{y}(t))\,, \tag{5.5}$$

and equipped with zero initial conditions

$$\mathbf{y}(t_0) = \mathbf{0}_N \quad , \quad \dot{\mathbf{y}}(t_0) = \mathbf{0}_N\,. \tag{5.6}$$

The $N_{\text{CD}} < N$ constraint equations arising from the Dirichlet boundary conditions are written in matrix form as

$$[B]^T\mathbf{y}(t) = \mathbf{0}_{N_{\text{CD}}} \quad , \quad t \in [t_0, T]\,, \tag{5.7}$$

where $[B]$ is a matrix in $\mathbb{M}_{N,N_{\text{CD}}}$ verifying $[B]^T[B] = [I_{N_{\text{CD}}}]$ and constructed using the N_{CD} zero Dirichlet conditions. In Eq. (5.4), the time-dependent applied force $\mathbf{f}(t)$ is directly constructed using Eq. (5.3). In Eq. (5.5), the nonlinear force $\mathbf{f}_{\text{NL}}(\mathbf{y}(t))$ induced by the elastic barriers is directly constructed using Eqs. (5.1) and (5.2). Note that the mass matrix $[M]$ belongs to $\mathbb{M}_N^+$, but the stiffness matrix $[K]$ is in $\mathbb{M}_N^{+0}$ and has a null space of dimension 6.

In general, the damping matrix $[\mathbb{D}]$ should also be in $\mathbb{M}_N^{+0}$ with the same null space as $[K]$. However, $[\mathbb{D}]$ is carefully constructed so that the modal damping rate obtained by projection of $[\mathbb{D}]$ onto the subspace spanned by the ROB constructed for this problem is effectively ξ_d. In addition, an adapted representation of $[\mathbb{D}]$ is carried out in order to avoid that $[\mathbb{D}]$ be a full matrix and that the computational complexity of a corresponding matrix-vector product grows with N^2. For the application considered here, a construction of $[\mathbb{D}]$ in $\mathbb{M}_N^+ \subset \mathbb{M}_N^{+0}$ is described in Appendix G.

In the frequency band of observation $\mathcal{B}_o = [0\,, \omega_o]$, the QoI (observation of the system) is the vector $\widehat{\mathbf{o}}(\omega) = (\widehat{o}_1(\omega), \ldots, \widehat{o}_{m_o}(\omega)) \in \mathbb{C}^{m_o}$ defined as follows:

- For plotting and analyzing the responses of the SROM, $m_{\mathbf{o}} = 2 \times 2 = 4$ dofs in the x_2 and x_3 directions of the 2 observation nodes Obs_{34} and Obs_{51} belonging to the aforementioned observation line (see Figure 12).
- For the identification of the hyperparameter $\boldsymbol{\alpha}$ of the SROB, $m_{\mathbf{o}} = 2 \times 61 = 122$ dofs in the x_2 and x_3 directions of all 61 nodes belonging to the observation line (see Figures 11 and 12).

For all ω in $\mathcal{B}_o$, the complex vector vector $\widehat{\mathbf{o}}(\omega)$ is written as

$$\widehat{\mathbf{o}}(\omega) = \mathbf{h}(-\omega^2\, \widehat{\mathbf{y}}(\omega)) \quad , \quad \omega \in \mathcal{B}_o\,, \tag{5.8}$$

where $\mathbf{h}$ is a linear mapping from $\mathbb{C}^N$ into $\mathbb{C}^{m_{\mathbf{o}}}$ which extracts the dofs from the vector $-\omega^2\, \widehat{\mathbf{y}}(\omega) \in \mathbb{C}^N$, and

$$\widehat{\mathbf{y}}(\omega) = \int_{t_0}^{T} e^{-i\omega\, t} \mathbf{y}(t)\, dt \quad , \quad \omega \in \mathcal{B}_o\,. \tag{5.9}$$

Note that $-\omega^2\, \widehat{\mathbf{y}}(\omega)$ would be the Fourier transform of $\{\ddot{\mathbf{y}}(t)\, t \in [t_0\,, T]\}$ if $\mathbf{y}(t_0) = \dot{\mathbf{y}}(t_0) = \mathbf{0}_N$ (which is the case, see Eq. (5.6)) and if $\mathbf{y}(T) = \dot{\mathbf{y}}(T) = \mathbf{0}_N$ (which is not exactly the case, because, as explained in Section 5.1, T corresponds to the time for which the dynamical system returns to the zero equilibrium with the relative error ε_d). Nevertheless, $\widehat{\mathbf{o}}(\omega)$ can be considered as a good approximation of the accelerations of the observations nodes in the x_2- and x_3 directions.

Also, Eqs. (5.4) and (5.5) are rewritten as $[M]\, \ddot{\mathbf{y}}(t) + [\mathbb{D}_d]\, \dot{\mathbf{y}}(t) + [K]\, \mathbf{y}(t) = \mathbf{f}(t) - \mathbf{f}_{\text{NL}}(\mathbf{y}(t)) - [\mathbb{D}_e]\, \dot{\mathbf{y}}(t)$, where the full $(N \times N)$ damping matrix is written as $[\mathbb{D}] = [\mathbb{D}_d] + [\mathbb{D}_e]$, $[\mathbb{D}_d]$ is the positive definite diagonal part of $[\mathbb{D}]$, and $[\mathbb{D}_e] = [\mathbb{D}] - [\mathbb{D}_d]$ is a full $(N \times N)$ damping matrix. The purpose of the latter matrix decomposition is to promote sparse computations (see Eq. (G.3) in Appendix G). The midpoint rule is applied to the time-integration of this equation using a fixed point method at each sampling time with the relative precision of 10^{-6}. In order to guarantee the convergence of the fixed point method, a local adaptive time-step is used.

The four subfigures of Figure 13 compare the graphs $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ predicted for this problem using the HDM described above and a linear variant obtained by supressing the nonlinear elastic barriers, for the x_2- and x_3-accelerations at Obs_{51} and Obs_{34}. Hence, these figures highlight the effects of the nonlinear elastic barriers on the response. In particular, they reveal an important energy transfer in the frequency band outside the main frequency band $[320\,, 620]\ Hz$ of the excitation.

5.3. Reduced-order model and performance

Let $\{\boldsymbol{\varphi}^1, \ldots, \boldsymbol{\varphi}^n\}$ be the first n elastic modes associated with the first n eigenfrequencies $0 < \omega_1 < \ldots < \omega_n$ of the linear undamped countpart model of the nonlinear damped dynamical system. These modes satisfy

$$[K]\, \boldsymbol{\varphi}^k = \lambda_k\, [M]\, \boldsymbol{\varphi}^k\,, \tag{5.10}$$

where $\lambda_1 = \omega_1^2, \ldots, \lambda_n = \omega_n^2$, and the constraint equation

$$[B]^T \boldsymbol{\varphi}^k = \mathbf{0}_{N_{\text{CD}}} \quad , \quad k = 1, \ldots, n\,. \tag{5.11}$$

Furthermore, the elastic modes satisfy the usual orthogonality properties

$$< [M]\, \boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= \delta_{kk'} \quad , \quad < [K]\, \boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= \lambda_k\, \delta_{kk'}\,. \tag{5.12}$$

For the nonlinear dynamical HDM, the ROB is chosen as $[V] = [\boldsymbol{\varphi}^1 \ldots \boldsymbol{\varphi}^n] \in \mathbb{M}_{N,n}$. This ROB satisfies the orthonormality property (2.5) and the constraint equation (2.4). Using this ROB and Eq. (5.12), the ROM associated with the HDM defined by Eqs. (5.4) to (5.7) can be written as

$$\mathbf{y}^{(n)}(t) = [V]\, \mathbf{q}(t) \quad , \quad t \in [t_0, T]\,, \tag{5.13}$$

$$\ddot{\mathbf{q}}(t) + [V]^T[\mathbb{D}][V]\, \dot{\mathbf{q}}(t) + [V]^T[K][V]\mathbf{q}(t) + [V]^T\mathbf{f}_{\text{NL}}([V]\mathbf{q}(t)) = [V]^T\mathbf{f}(t)\ ,\ t \in]t_0, T]\,, \tag{5.14}$$

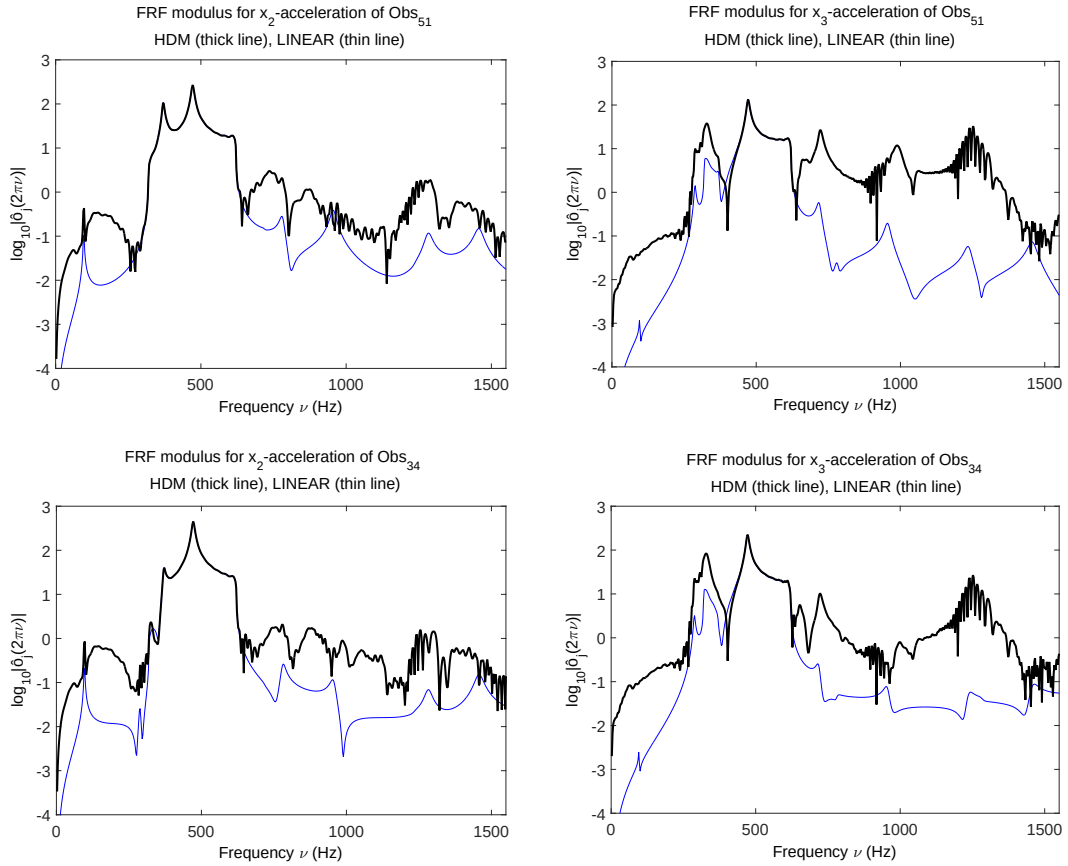


Figure 13. Graphs of $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ computed using the HDM (thick lines) and a linear counterpart (thin lines) for the x_2- and x_3-accelerations at Obs_{51} (top left and top right) and at Obs_{34} (bottom left and bottom right).

with the initial conditions

$$\mathbf{q}(0) = \mathbf{0}_n \quad , \quad \dot{\mathbf{q}}(0) = \mathbf{0}_n \, .$$

The approximation $\widehat{\mathbf{o}}^{(n)}(\omega)$ of the QoI $\widehat{\mathbf{o}}(\omega)$ defined by Eqs. (5.8) and (5.9) can be written as

$$\widehat{\mathbf{o}}^{(n)}(\omega) = \mathbf{h}(-\omega^2 \, \widehat{\mathbf{y}}^{(n)}(\omega)) \quad , \quad \omega \in \mathcal{B}_o \, , \tag{5.15}$$

$$\widehat{\mathbf{y}}^{(n)}(\omega) = \int_{t_0}^{T} e^{-i\omega\, t} \mathbf{y}^{(n)}(t) \, dt \quad , \quad \omega \in \mathcal{B}_o \, . \tag{5.16}$$

To solve the above reduced governing equations, the algorithm described in Section 5.2 is used after it is adapted to Eq. (5.14), and this equation is rewritten, using Eqs. (G.2) and (5.12), as $\ddot{\mathbf{q}}(t) + 2\,\xi_d\,[\lambda^{(n)}]^{1/2}\,\dot{\mathbf{q}}(t) + [\lambda^{(n)}]\,\mathbf{q}(t) = [V]^T\mathbf{f}(t) - [V]^T\mathbf{f}_{\text{NL}}([V]\,\mathbf{q}(t))$, where $[\lambda^{(n)}]$ is the positive definite diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$.

Taking into account the fact that a significative difference between the responses delivered by the HDM and ROM must be generated in order to demonstrate the capability of the proposed nonparametric stochastic method to account for this type of modeling errors, a good compromise between computational cost and ROM accuracy leads to choosing $n = 20$ as the dimension of the ROM. In this case, $\omega_1 = 2\,\pi \times 96.69 \; rad/s$, $\omega_4 = 2\,\pi \times 472 \; rad/s$, $\omega_5 = 2\,\pi \times 720 \; rad/s$, $\omega_{11} = 2\,\pi \times 1474 \, rad/s$, $\omega_{12} = 2\,\pi \times 1754 \, rad/s$, and $\omega_{20} = 2\,\pi \times 2936 \, rad/s$. Consequently, there are 4 elastic modes in the frequency band $[0\,,620]\; Hz$ containing the main part of the excitation, 11 elastic

modes in the frequency band of observation $[0\,,1550]\ Hz$, and 9 elastic modes in the frequency band $[1550\,,3100]\ Hz$. In Figure 14, the four subfigures compare the graphs $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ computed using the HDM with the graphs $\nu \mapsto \log_{10}(|\widehat{o}_j^{(n)}(2\pi\nu)|)$ computed using the chosen ROM, for the x_2- and x_3-accelerations at Obs_{51} and Obs_{34}. They show that the differences between the HDM and ROM predictions are very small in the frequency band $[320\,,620]\ Hz$ of the excitation, but significant outside this frequency band (by design, such differences can be reduced by increasing the dimension n of the ROM).

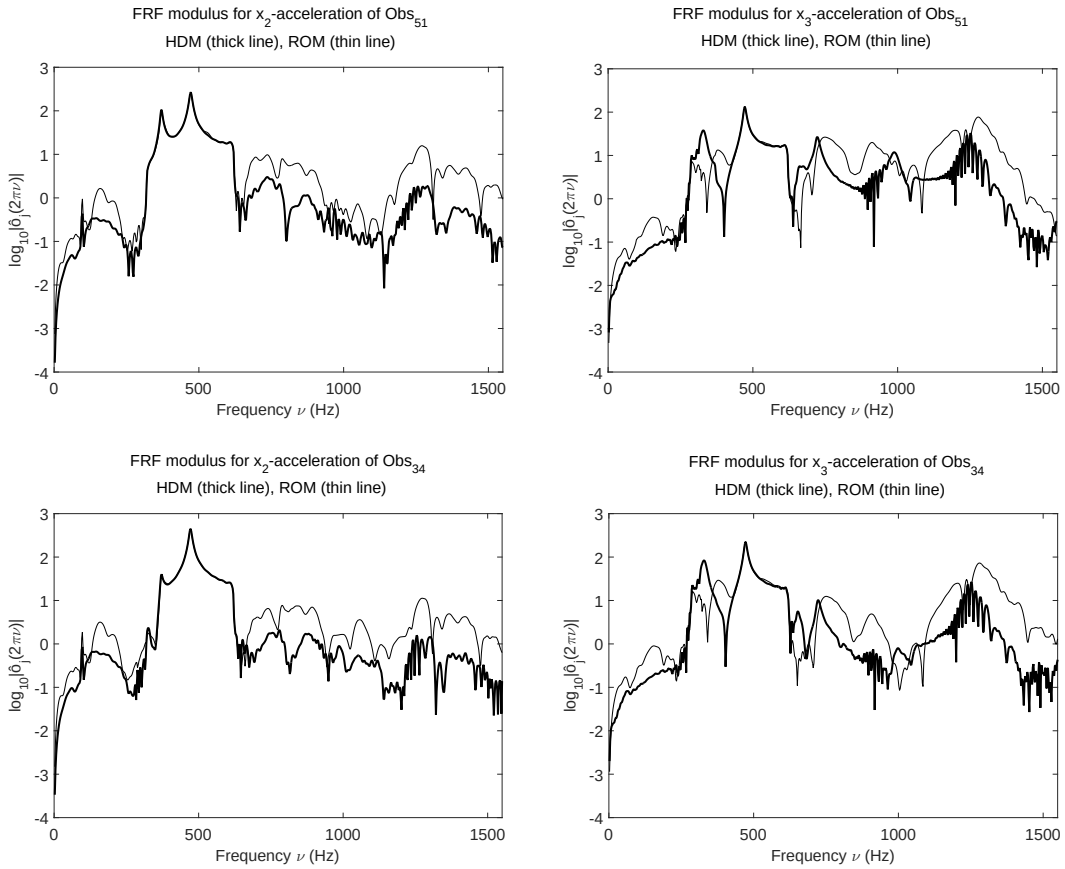


Figure 14. Graphs of $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ computed using the HDM (thick lines) and ROM (thin lines) for the x_2- and x_3-accelerations at Obs_{51} (top left and top right) and Obs_{34} (bottom left and bottom right).

5.4. Stochastic reduced-order model and performance

From Eqs. (2.13) to (2.16), it follows that the SROM associated with Eqs. (5.13) to (5.16) is obtained by substituting $[V]$ with the random matrix $[\mathbf{W}]$. Consequently, $\mathbf{y}^{(n)}$, $\mathbf{q}$, and $\mathbf{o}^{(n)}$ become the random vectors $\mathbf{Y}^{(n)}$, $\mathbf{Q}$, and $\widehat{\mathbf{O}}^n$, and the SROM becomes

$$\mathbf{Y}^{(n)}(t) = [\mathbf{W}]\,\mathbf{Q}(t) \quad , \quad t \in [t_0, T]\,,$$

$$\ddot{\mathbf{Q}}(t) + [\mathbf{W}]^T[\mathbb{D}][\mathbf{W}]\,\dot{\mathbf{Q}}(t) + [\mathbf{W}]^T[K][\mathbf{W}]\mathbf{Q}(t) + [\mathbf{W}]^T\mathbf{f}_{\text{NL}}([\mathbf{W}]\mathbf{Q}(t)) = [\mathbf{W}]^T\mathbf{f}(t)\,,\ t \in]t_0, T]\,, \quad (5.17)$$

with the initial conditions

$$\mathbf{Q}(0) = \mathbf{0}_n \quad , \quad \dot{\mathbf{Q}}(0) = \mathbf{0}_n \quad , a.s \quad .$$

Let $\widehat{\mathbf{O}}^{(n)}(\omega;\boldsymbol{\alpha})$ be the random QoI with values in $\mathbb{C}^{m_o}$ and with $m_o = 122$. This random QoI depends on the hyperparameter $\boldsymbol{\alpha}$ and is such that

$$\widehat{\mathbf{O}}^{(n)}(\omega;\boldsymbol{\alpha}) = \mathbf{h}(-\omega^2\,\widehat{\mathbf{Y}}^{(n)}(\omega;\boldsymbol{\alpha})) \quad , \quad \omega \in \mathcal{B}_o\,, \tag{5.18}$$

$$\widehat{\mathbf{Y}}^{(n)}(\omega;\boldsymbol{\alpha}) = \int_{t_0}^{T} e^{-i\omega\,t}\mathbf{Y}^{(n)}(t;\boldsymbol{\alpha})\,dt \quad , \quad \omega \in \mathcal{B}_o\,.$$

In Eq. (5.17), the $\mathbb{M}_n^+$-valued random matrix $[\mathbf{W}]^T[\mathbb{D}][\mathbf{W}]$ is constructed using Eq. (G.4), which shows that the $(N \times N)$ full matrix $[\mathbb{D}]$ is not assembled.

To identify the hyperparameter $\boldsymbol{\alpha} = (s, \beta, \sigma)$ defined in Section 3.3.3, of length $m_{\boldsymbol{\alpha}} = 2 + n(n+1)/2 = 212$, and belonging to the admissible set $\mathcal{C}_{\boldsymbol{\alpha}} = \mathbb{R}^+ \times [0.01\,,0.1] \times \mathbb{M}_n^{\mathrm{u}}$, the cost function $J(\boldsymbol{\alpha})$ defined by Eq. (H.3) of Appendix H, where $w_J = 0.9$ and $\gamma = 0.3$, is minimized. Specifically, the optimization problem

$$\boldsymbol{\alpha}^{\mathrm{opt}} = (s^{\mathrm{opt}}, \beta^{\mathrm{opt}}, [\sigma^{\mathrm{opt}}]) = \min_{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\alpha}}} J(\boldsymbol{\alpha})\,, \tag{5.19}$$

is solved using the algorithm described in Appendix F. The Monte Carlo solver is used with $1\,000$ independent realizations (mean-square convergence reached).

Figure 15 (left) displays the graph of the function $\beta \mapsto J(\boldsymbol{\alpha}(\beta))$, where $\boldsymbol{\alpha}(\beta) = (s^{(1)}, \beta, [\sigma^{(2)}])$ (used in Stage 3 of the optimization algorithm described in Appendix F) and Figure 2 (right) displays the graph of the diagonal entries of the function $k \mapsto [c_n^{\mathrm{opt}}]_{kk}$, where $[c_n^{\mathrm{opt}}] = [\sigma^{\mathrm{opt}}]^T[\sigma^{\mathrm{opt}}]$ corresponds to the optimal value $[\sigma^{\mathrm{opt}}]$ of $[\sigma]$ (see Eq. (3.16)). For the sake of limiting the number of figures, the extra diagonal entries of $[c_n^{\mathrm{opt}}]$ are not displayed. The optimal values of s and β are found to be $s^{\mathrm{opt}} = 0.0103$ and $\beta^{\mathrm{opt}} = 0.0181$.

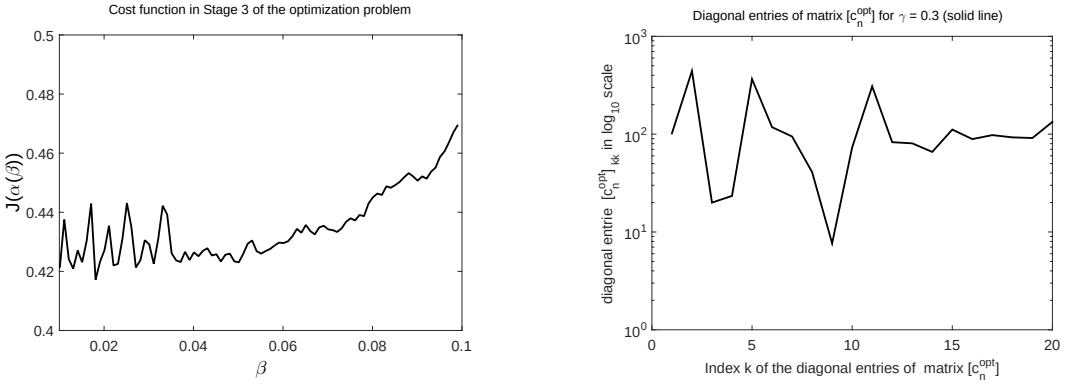


Figure 15. Graph of $\beta \mapsto J(\boldsymbol{\alpha}(\beta))$, where $\boldsymbol{\alpha}(\beta) = (s^{(1)}, \beta, [\sigma^{(2)}])$ (left), and of $k \mapsto [c^{\mathrm{opt}}]_{kk}$ (right, using the $\log_{10}$ scale) for $\gamma = 0.3$ (solid line).

The four subfigures of Figure 16 summarize the results obtained using the SROM. These subfigures pertain to the x_2- and x_3-accelerations at Obs_{51} and Obs_{34}. Each of them displays the graph of $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ computed using the HDM (the target for the mean), that of $\nu \mapsto \log_{10}(|\widehat{o}_j^{(n)}(2\pi\nu)|)$ computed using the nonlinear ROM, and the confidence region, with a probability $p_c = 0.98$, of the frequency sampled stochastic process $\nu \mapsto \log_{10}(|\widehat{O}_j^{(n)}(2\pi\nu)|)$ constructed using the SROM. The upper envelope of the confidence region corresponds to the quantile for the probability p_c, and the lower envelope to the quantile for the probability $1 - p_c$. The reader can observe that the obtained results are very good, except for the x_3-accelerations in the small part $[270\,,438]\ Hz$ of the frequency band $[0\,,1550]\ Hz$ of the analysis. This prediction can be improved by using a more sophisticated optimization algorithm than that used here for

the identification of the hyperparameters. Nevertheless, it can be seen that the SROM allows the generation of a confidence region which is not centered around the responses computed using the ROM, but is approximatively well centered around the responses computed using the HDM. This demonstrates the capability of the proposed method for accounting for second-order contributions.

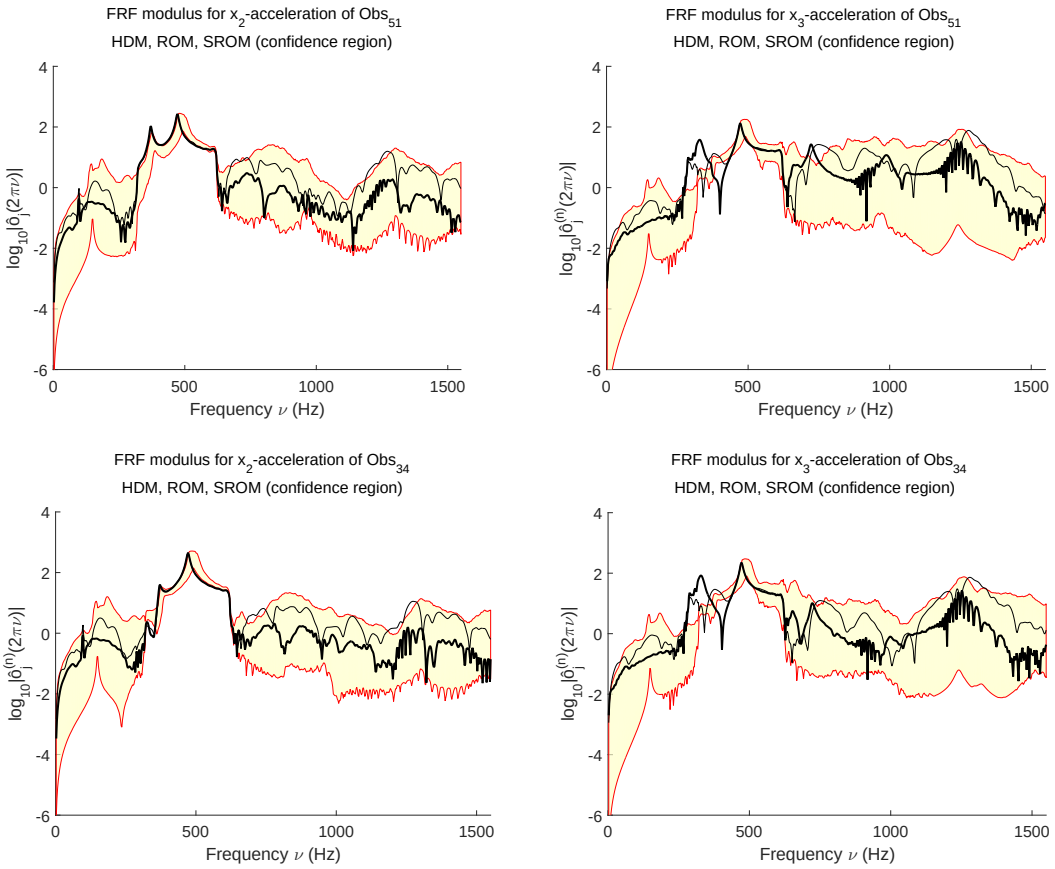


Figure 16. Graphs of $\nu \mapsto \log_{10}(|\widehat{o}_j(2\pi\nu)|)$ computed using the HDM (thick lines) and ROM (thin lines), and confidence region for $p_c = 0.98$ (yellow region with red upper and red lower envelopes) of $\nu \mapsto \log_{10}(|\widehat{O}_j^{(n)}(2\pi\nu)|)$ computed using the SROM, for the x_2- and x_3-accelerations at Obs_{51} (top left and top right) and Obs_{34} (bottom left and bottom right).

6. APPLICATION TO A PARAMETERIZED NONLINEAR COMPUTATIONAL STRUCTURAL DYNAMICS PROBLEM

Whereas Sections 4 and 5 have focused on two nonparametric problems, this section illustrates the capabilities of the proposed nonparametric probabilistic method in the case of a $\boldsymbol{\mu}$-parametric HDM. In particular, the objective of this section is to analyze the performance of a $\boldsymbol{\mu}$-parametric SROM constructed as proposed in this paper at predicting a confidence region that contains the solution of the $\boldsymbol{\mu}$-parametric HDM, for a quality-assessment value $\boldsymbol{\mu}_{\text{qa}}$ of $\boldsymbol{\mu}$ which does not belong to the subset of training points $\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{m_\mu}\}$ of $\mathcal{C}_{\boldsymbol{\mu}}$. To this effect, the reader is reminded that the training points are used for constructing the ROB $[V]$, identifying $\boldsymbol{\alpha}^{\text{opt}}$, and therefore, constructing the SROB $\mathbf{W}$. For this purpose, the simple example presented in Sections 4.1 to 4.6 is reused as described next.

6.1. Description of the μ-parametric HDM

Consider again the linear static problem described in Section 4 with $N = 1\,000$, and such that Eq. (2.1) is rewritten as

$$[K]\,\mathbf{y}(\mu) = \mathbf{f}(\mu)\,. \tag{6.1}$$

The vector $\mathbf{y}(\mu) = (y_1(\mu), \ldots, y_N(\mu))$ belonging to $\mathbb{R}^N$ represents the dimensionless displacement. The matrix $[K]$ belonging to $\mathbb{M}_N^+$ is the stiffness matrix of the problem. The vector $\mathbf{f}(\mu) = (f_1(\mu), \ldots, f_N(\mu))$ belonging to $\mathbb{R}^N$ is the dimensionless force with $f_1(\mu) = f_N(\mu) = 0$; it depends on the real-valued parameter μ belonging to $\mathcal{C}_{\boldsymbol{\mu}} = [\,\mu_{\min}\,,\mu_{\max}\,] \subset \mathbb{R}$, with with $\mu_{\min} = 0.5$ and $\mu_{\max} = 1.1$. The matrix $[K]$ and vector $\mathbf{f}(\mu)$ are defined/constructed as described in Appendix I. Eq. (6.1) has a unique solution which satisfies $y_1(\mu) = y_N(\mu) = 0$ (Dirichlet conditions), and consequently, $\mathbf{y}(\mu)$ satisfies the following equation

$$[B]^T\mathbf{y}(\mu) = \mathbf{0}_{N_{\text{CD}}} \quad , \quad [B]^T[B] = [I_N]\,, \tag{6.2}$$

where $[B]$ belongs to $\mathbb{M}_{N,N_{\text{CD}}}$ with $N_{\text{CD}} = 2$ and is explicitly constructed as described in Appendix I. For this example, the QoI (see Eq. (2.3)) is defined as $\mathbf{o}^D(\mu) = (o_1^D(\mu), \ldots, o_N^D(\mu))$ — that is, the vector in $\mathbb{R}^{m_{\mathbf{o}}}$, with $m_{\mathbf{o}} = N$, which satisfies

$$\mathbf{o}^D(\mu) = \mathbf{y}(\mu)\,,$$

where $\mathbf{y}(\mu)$ is the unique solution of Eq. (6.1).

6.2. Construction of the μ-parametric ROM

To construct the ROM, the parameter space $\mathcal{C}_\mu$ is sampled at the following $m_\mu = 3$ points

$$\mu_1 = 0.5 \quad , \quad \mu_2 = 0.7 \quad , \quad \mu_3 = 1.1\,.$$

For each sampled value μ_i, $i = 1, 2, 3$, the solution of Eq. (6.1) under the constraint (6.2) is denoted by $\mathbf{y}(\mu_i)$. The singular values of the resulting matrix $[\mathbf{y}(\mu_1)\ \mathbf{y}(\mu_2)\ \mathbf{y}(\mu_3)] \in \mathbb{M}_{N,3}$ are $s_1 = 5.9 \times 10^{-2}$, $s_2 = 6.1 \times 10^{-3}$, and $s_3 = 1.2 \times 10^{-15}$. In view of this, and in order to construct a ROM with modeling errors, a ROB $[V] \in \mathbb{M}_{N,1}$ ($n = 1$) is constructed using the right singular vector associated with the largest singular value. Hence, this ROB satisfies

$$[V]^T\,[V] = [I_1] \quad , \quad [B]^T\,[V] = [0_{2,1}]\,,$$

which is consistent with the choice $[M] = [I_N]$ for static problems. Then, the μ-parametric ROM defined by Eqs. (2.6) and (2.7) is rewritten as

$$\mathbf{y}^{(n)}(\mu) = [V]\,\mathbf{q}(\mu)\,, \tag{6.3}$$

$$[V]^T[K]\,[V]\,\mathbf{q}(\mu) = [V]^T\mathbf{f}(\mu)\,. \tag{6.4}$$

The corresponding approximation $\mathbf{o}^{(D,n)}(\mu)$ of the QoI $\mathbf{o}^{(D,n)}(\mu) = (o_1^{(D,n)}(\mu), \ldots, o_N^{(D,n)}(\mu))$ is written (see Eq. (2.9)) as

$$\mathbf{o}^{(D,n)}(\mu) = \mathbf{y}^{(n)}(\mu)\,. \tag{6.5}$$

6.3. Performance of the μ-parametric ROM

To assess its accuracy for parametric computations, the constructed ROM is applied to predict the solution of problem (6.1) at the unsampled parameter value $\mu_{\text{qa}} = 0.6$. Figure 17 displays the graph of $j \mapsto o_j^{(D,n)}(\mu)$ computed using the constructed ROM, which reveals some expected inaccuracies of the μ-parametric ROM.

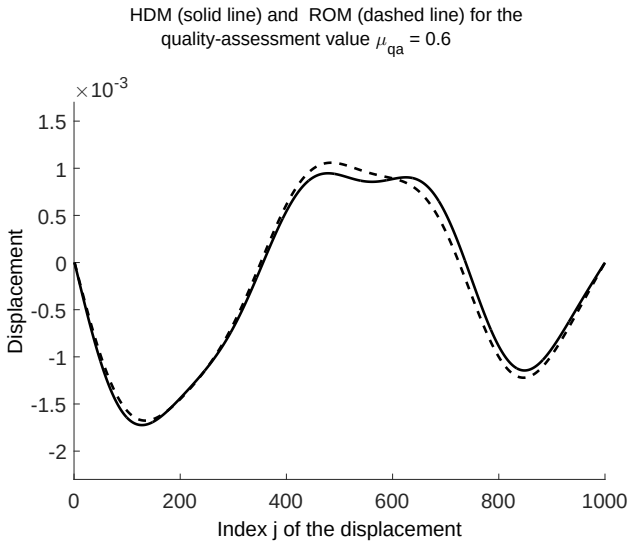


Figure 17. Graph of $j \mapsto o_j^D(\mu_{\text{qa}})$ computed using the HDM (solid line) and counterpart graph of $j \mapsto o_j^{(D,n)}(\mu_{\text{qa}})$ computed using the ROM (dashed line).

6.4. Construction of the μ-parametric SROM

From Eqs. (2.13) to (2.16), it follows that the μ-parametric SROM associated with Eqs. (6.3) to (6.5) is obtained by substituting $[V]$ with the random matrix $[\mathbf{W}]$. Consequently, $\mathbf{y}^{(n)}(\mu)$, $\mathbf{q}(\mu)$, and $\mathbf{o}^{(n)}(\mu)$, become the random vectors $\mathbf{Y}^{(n)}(\mu,\boldsymbol{\alpha})$, $\mathbf{Q}(\mu,\boldsymbol{\alpha})$, and $\mathbf{O}^{(D,n)}(\mu,\boldsymbol{\alpha})$, all of which depend on the parameter μ and the hyperparameter $\boldsymbol{\alpha}$. Hence, the μ-parametric SROM can be written for this problem as

$$\mathbf{Y}^{(n)}(\mu,\boldsymbol{\alpha}) = [\mathbf{W}]\,\mathbf{Q}(\mu,\boldsymbol{\alpha})\,, \tag{6.6}$$

$$[\mathbf{W}]^T[K]\,[\mathbf{W}]\,\mathbf{Q}(\mu,\boldsymbol{\alpha}) = [\mathbf{W}]^T\,\mathbf{f}(\mu)\,, \tag{6.7}$$

$$\mathbf{O}^{(D,n)}(\mu,\boldsymbol{\alpha}) = \mathbf{Y}^{(n)}(\mu,\boldsymbol{\alpha})\,. \tag{6.8}$$

6.5. Identification of the hyperparameter of the SROB

The hyperparameter $\boldsymbol{\alpha}$ defined in Section 3.3.3 is written as $\boldsymbol{\alpha} = (s,\beta,\sigma)$. In this case ($n = 1$), its length is $m_{\boldsymbol{\alpha}} = 2 + n(n+1)/2 = 3$. It is identified using the nonlinear LS method presented in Section 2, after adaptation to the present example as described below.

Specifically, the cost function (2.17) is considered with

$$J_{\text{mean}}\,(\boldsymbol{\alpha}) = \frac{\sum_{i=1}^{m_\mu} \|\mathbf{o}^D(\mu_i) - E\{\mathbf{O}^{(D,n)}(\boldsymbol{\alpha},\mu_i)\}\|^2}{\sum_{i=1}^{m_\mu} \|\mathbf{o}^D(\mu_i)\|^2}\,,$$

$$J_{\text{std}}\,(\boldsymbol{\alpha}) = \frac{\sum_{i=1}^{m_\mu} \|\mathbf{v}^D(\mu_i) - \mathbf{v}^{(D,n)}(\mu_i,\boldsymbol{\alpha})\|^2}{\sum_{i=1}^{m_\mu} \|\mathbf{v}^D(\mu_i)\|^2}\,, \tag{6.9}$$

where $\mathbf{v}^D(\mu_i) = (v_1^D(\mu_i),\ldots,v_{m_{\mathbf{o}}}^D(\mu_i))$ is such that

$$v_j^D(\mu_i) = \gamma\,|o_j^D(\mu_i) - o_j^{(D,n)}(\mu_i)| \quad , \quad j = 1,\ldots m_{\mathbf{o}}\,,$$

and $\gamma > 0$. In Eq. (6.9), the vector $\mathbf{v}^{(D,n)}(\mu_i,\boldsymbol{\alpha})$ whose components are $(v_1^{(D,n)}(\mu_i,\boldsymbol{\alpha}),\ldots,v_{m_{\mathbf{o}}}^{(D,n)}(\mu_i,\boldsymbol{\alpha}))$ is such that

$$v_j^{(D,n)}(\mu_i,\boldsymbol{\alpha}) = \{\,E\{O_j^{(D,n)}(\mu_i,\boldsymbol{\alpha})^2\} - (E\{O_j^{(D,n)}(\mu_i,\boldsymbol{\alpha})\})^2\}\,\}^{1/2}\,, \quad j = 1,\ldots m_{\mathbf{o}}\,. \tag{6.10}$$

The optimization problem is

$$\boldsymbol{\alpha}^{\text{opt}} = (s^{\text{opt}}, \beta^{\text{opt}}, \sigma^{\text{opt}}) = \min_{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\alpha}}} J(\boldsymbol{\alpha})\,. \tag{6.11}$$

The parameters γ and w_J introduced in Section 6.5 are fixed to $\gamma = 0.8$ and $w_J = 0.5$. The above optimization problem is solved by using only stage 1 of the algorithm presented in Appendix F (for $n = 1$, the matrix $[\sigma]$ becomes a scalar σ and consequently, the hyperparameter $\boldsymbol{\alpha}$ has only 3 components, namely, are s, β, and σ). The admissible set $\mathcal{C}_{\boldsymbol{\alpha}}$ is defined by Eq. (3.25) with $\varepsilon_0 = 0.01$, $\beta_d = 0.15$, $\beta_u = 0.30$, and $\sigma_u = 20$. For the interior-point algorithm, the initial values are selected as follows: $s_0 = 0.017$, $\beta_0 = 0.2$, and $\sigma_0 = 1$. The optimal value of s, β, and σ are found to be $s^{\text{opt}} = 0.0869$, $\beta^{\text{opt}} = 0.21$, and $\sigma^{\text{opt}} = 0.74$.

6.6. Performance of the μ-parametric SROM at an unsampled parameter value

The stochastic solution of Eqs. (6.6) to (6.8), where $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{\text{opt}}$, is computed using the Monte Carlo method with $1\,000$ independent realizations (for which the mean-square convergence is reached with a reasonable accuracy).

For the unsampled parameter value $\mu = \mu_{qa} = 0.6$, Figure 18 (left) displays the graph of $j \mapsto o_j^{(D,n)}(\mu_{\text{qa}})$ of the displacement field computed with the ROM, and that of $j \mapsto E\{O_j^{(D,n)}(\mu_{\text{qa}}, \boldsymbol{\alpha}^{\text{opt}})\}$ of the mean function of the random displacement field computed using the SROM. The reader can observe that the mean function computed using the SROM is relatively close to that computed using the ROM. Figure 18 (right) displays the graph of $j \mapsto v_j^{(D,n)}(\mu_{\text{qa}}, \boldsymbol{\alpha}^{\text{opt}})$ of the standard deviation (see Eq. (6.10)) of the random displacement field computed using the SROM.

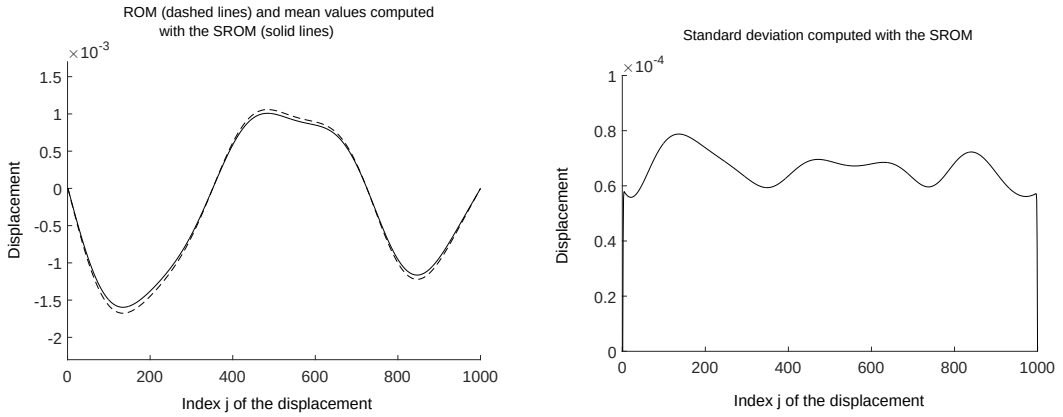


Figure 18. Left figure: graph of $j \mapsto o_j^{(D,n)}(\mu)$ computed using the ROM (dashed line) and that of $j \mapsto E\{O_j^{(D,n)}(\mu_{\text{qa}}, \boldsymbol{\alpha}^{\text{opt}})\}$ of the mean function computed using the SROM (solid line) — Right figure: Graph of $j \mapsto v_j^{(D,n)}(\mu_{,\text{qa}}\, \boldsymbol{\alpha}^{\text{opt}})$ of the standard deviation computed using the SROM (solid line).

Finally, Figure 19 displays the graph of $j \mapsto o_j^D(\mu_{\text{qa}})$ computed using the HDM, that of $j \mapsto o_j^{(D,n)}(\mu_{\text{qa}})$ computed using the ROM, and the confidence region for $p_c = 0.98$ of $j \mapsto O_j^{(D,n)}(\mu_{\text{qa}}, \boldsymbol{\alpha}^{\text{opt}})$ computed using the SROM. It demonstrates a good quality of the μ-parametric SROM

7. CONCLUSIONS

A nonparametric probabilistic approach for quantifying uncertainties in a given linear or nonlinear High-Dimensional Model (HDM) for which a projection-based Reduced-Order Model (ROM) can

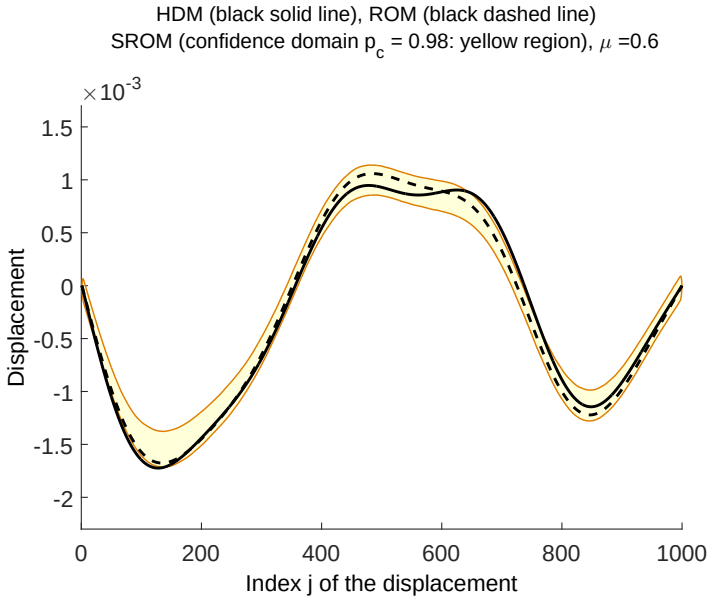


Figure 19. Graph of $j \mapsto o_j^D(\mu_{\text{qa}})$ computed using the HDM (black solid line), that of $j \mapsto o_j^{(D,n)}(\mu_{\text{qa}})$ computed using the ROM (black dashed line), and for $p_c = 0.98$, confidence region (yellow region with red upper and lower envelopes) of $j \mapsto O_j^{(D,n)}(\mu_{\text{qa}}, \boldsymbol{\alpha}^{\text{opt}})$ computed using the SROM.

be constructed is presented. The underlying probabilistic model is implemented directly in the ROM. The associated stochastic model is characterized by a small number of hyperparameters whose identification via the solution of a statistical inverse problem is computationally feasible. The cost function describing this inverse problem is formulated with respect to a given target related to given observations, which allows for specifying the level of uncertainties induced by various model form uncertainties introduced during the construction of both the HDM and ROM. While this approach for modeling uncertainties is developed for large-scale nonlinear computational models, it is demonstrated in this paper using simple academic examples. These are nevertheless representative of realistic problems, and have the advantage of being easy to reproduce by the interested reader. All obtained numerical results highlight the potential of the proposed nonparametric stochastic model of model uncertainties to control the statistical fluctuations of the random eigenvalues of a linear operator, the statistical mean values and fluctuations of the random solution of a linear elliptic problem, and those of a nonlinear dynamical problem.

ACKNOWLEDGEMENTS

The second author acknowledges partial support by the Army Research Laboratory through the Army High Performance Computing Research Center under Cooperative Agreement W911NF-07-2-0027, partial support by The Boeing Company under Contract Sponsor Ref. 45047, and partial support by DARPA under the Enabling Quantification of Uncertainty in Physical Systems (EQUiPS) program. This document does not necessarily reflect the position of these institutions, and no official endorsement should be inferred.

APPENDIX A: PROOF OF EQ. (3.27)

(i) The proof of equation $[\underline{W}(s,\beta,\sigma)] = [V]\, E\{[H_s(\mathbf{Z})]\}$ is the following. Equation (3.19) yields $[\underline{W}(s,\beta,\sigma)] = [V]\, E\{[H_s(\mathbf{Z})]\} + sE\{[\mathbf{Z}]\,[H_s(\mathbf{Z})]\}$. We then have to prove that $E\{[\mathbf{Z}]\,[H_s(\mathbf{Z})]\} = [0_{n,n}]$. From Eqs. (3.21) to (3.24), random matrix $[\mathbf{Z}]$ can be written as $[\mathbf{Z}] = [z(\mathbf{G})]$ in which the mapping $[G] \mapsto [z(G)]$ is defined on $\mathbb{M}_{N,n}$. We have

$$E\{[\mathbf{Z}]\,[H_s(\mathbf{Z})]\} = \int_{\mathbb{M}_{N,n}} [z(G)]\,[H_s(z(G))]\, p_{[\mathbf{G}]}([G])\, d[G]\,. \tag{A.1}$$

Changing $[G]$ in $-[G]$ in the right-hand side of Eq. (A.1), and since $p_{[\mathbf{G}]}(-[G]) = p_{[\mathbf{G}]}([G])$ (see Eq. (D.59)), $[z(-G)] = -[z(G)]$, $[H_s(z(-G))] = [H_s(-z(G))] = [H_s(z(G))]$, Eq. (A.1) yields

$$E\{[\mathbf{Z}]\,[H_s(\mathbf{Z})]\} = -\int_{\mathbb{M}_{N,n}} [z(G)]\,[H_s(z(G))]\, p_{[\mathbf{G}]}([G])\, d[G]\,. \tag{A.2}$$

Adding Eq. (A.1) with Eq. (A.2) yields $E\{[\mathbf{Z}]\,[H_s(\mathbf{Z})]\} = [0_{n,n}]$.

(ii) The proof of equation $[B]^T[\underline{W}(s,\beta,\sigma)] = [0_{N_{\text{CD}},n}]$ is the following. Since $[B]^T[\mathbf{W}] = [0_{N_{\text{CD}},n}]$ almost surely, and since $E\{[\mathbf{W}]\}$ exists, it can be concluded that $E\{[B]^T[\mathbf{W}]\} = [0_{N_{\text{CD}},n}]$, and consequently, $[B]^T\, E\{[\mathbf{W}]\} = [0_{N_{\text{CD}},n}]$.

APPENDIX B: PROOFS OF EQ. (3.28) AND EQ. (3.29)

Since $E\{\|\mathbf{W} - \underline{W}(s,\beta,\sigma)\|_M^2\} = E\{\text{tr}\{[\mathbf{W}]^T[M]\,[\mathbf{W}]\}\} - \text{tr}\{[\underline{W}(s,\beta,\sigma)]^T[M]\,[\underline{W}(s,\beta,\sigma)]\}$, from Eq. (2.12), Eq. (2.11) and (3.26), it can be deduced that

$$\text{var}_{\mathbf{w}}^M(s,\beta,\sigma) = n - \|\underline{W}(s,\beta,\sigma)\|_M^2 \quad , \quad \forall\, s \geq 0\ ,\ \beta > 0\ ,\ [\sigma] \in \mathbb{M}_n^u\,. \tag{B.1}$$

Since $\text{var}_{\mathbf{w}}^M(s,\beta,\sigma) \geq 0$, Eq. (B.1) yields $\|\underline{W}(s,\beta,\sigma)\|_M^2 \leq n$, which proves the first equation in Eq. (3.28). For $s = 0$, Eq. (B.1) yields $\text{var}_{\mathbf{w}}^M(0,\beta,\sigma) = n - \|\underline{W}(0,\beta,\sigma)\|_M^2$ and since $[\underline{W}(0,\beta,\sigma)]^T[M]\,[\underline{W}(0,\beta,\sigma)] = [I_n]$, it can be deduced that $\text{var}_{\mathbf{w}}^M(0,\beta,\sigma) = 0$, which proves the first equation in Eq. (3.29). From Eqs. (3.27) and (3.1), we can write

$$\|\underline{W}(s,\beta,\sigma)\|_M^2 = \|E\{H_s(\mathbf{Z})\}\|_F^2\,. \tag{B.2}$$

It can be seen that,

$$s \to +\infty \quad , \quad E\{[H_s(\mathbf{Z})]\} \sim \frac{1}{s}\,[J(\beta,\sigma)] \quad , \quad \forall\, \beta > 0\ ,\ [\sigma] \in \mathbb{M}_n^u\,, \tag{B.3}$$

in which $[J(\beta,\sigma)] = E\{([\mathbf{Z}]^T[M]\,[\mathbf{Z}])^{-1/2}\}$ is a matrix in $\mathbb{M}_n^+$ depending on β and $[\sigma]$ (but independent of s) such that $\|J(\beta,\sigma)\|_F < +\infty$. Using Eqs. (B.2) and (B.3) yields

$$s \to +\infty \quad , \quad \|\underline{W}(s,\beta,\sigma)\|_M^2 \sim \frac{1}{s^2}\,\|J(\beta,\sigma)\}\|_F^2 \quad , \quad \forall\, \beta > 0\ ,\ [\sigma] \in \mathbb{M}_n^u\,. \tag{B.4}$$

Eqs. (B.1) and (B.4) yield

$$s \to +\infty \quad , \quad \text{var}_{\mathbf{w}}^M(s,\beta,\sigma) \sim n - \frac{1}{s^2}\,\|J(\beta,\sigma)\}\|_F^2 \quad , \quad \forall\, \beta > 0\ ,\ [\sigma] \in \mathbb{M}_n^u\,. \tag{B.5}$$

From Eq. (B.5), it can be deduced the second equation in Eq. (3.29) and that the mapping $s \mapsto \text{var}_{\mathbf{w}}^M(s,\beta,\sigma)$ is an increasing function that is asymptotic to n.

APPENDIX C: PROOF OF EQ. (3.30)

For all $\beta > 0$ and for all $[\sigma]$ in $\mathbb{M}_n^u$, Eqs. (3.19) and (3.20) yield

$$\lim_{s\to+\infty} [\mathbf{W}(s,\beta,\sigma)] = [\mathbf{Z}]([\mathbf{Z}]^T[M]\,[\mathbf{Z}])^{-1/2} \quad a.s. \tag{C.1}$$

We have $\ll [\mathbf{W}(s,\beta,\sigma)]\,,[V] \gg_M = \mathrm{tr}\{[\mathbf{W}(s,\beta,\sigma)]^T[M]\,[V]\} = \mathrm{tr}\{[V]^T[M]\,[\mathbf{W}(s,\beta,\sigma)]\}$ and consequently, from Eq. (C.1), it can be deduced that $\lim_{s\to+\infty} \ll [\mathbf{W}(s,\beta,\sigma)]\,,[V] \gg_M = \mathrm{tr}\{[a]\,[\mathbf{S}]\}$, in which $[a] = [V]^T[M]\,[\mathbf{Z}]$ is a random skew-symmetric matrix (see Section 3.1) and where $[\mathbf{S}] = ([\mathbf{Z}]^T[M]\,[\mathbf{Z}])^{-1/2}$ is a random symmetric matrix. Since the trace of the product of a skew-symmetric matrix with a symmetric one is zero, we obtain $\lim_{s\to+\infty} \ll [\mathbf{W}(s,\beta,\sigma)]\,,[V] \gg_M = 0$.

APPENDIX D: CONTINUOUS INTERPRETATION OF THE SROB AND CONSTRUCTION OF THE RANDOM GENERATOR OF RANDOM MATRIX $[\mathbf{G}(\beta)]$

In this appendix, we present a continuous interpretation of the SROB constructed in Section 3.3 and, in Section D.5, we detail the generator of independent realizations for random matrix $[\mathbf{G}(\beta)]$.

D.1. Computational SROB to be deduced from a continuous SROB

The problem is to give the continuous version of the following computational SROB $[\mathbf{W}]$ that is defined by the following equations,

$$[\mathbf{W}] = R_{s,V}([\mathbf{Z}]) = ([V] + s\,[\mathbf{Z}])\,[H_s(\mathbf{Z})]\,, \tag{D.1}$$

$$[H_s(\mathbf{Z})] = ([I_n] + s^2[\mathbf{Z}]^T[M]\,[\mathbf{Z}])^{-1/2}\,, \tag{D.2}$$

$$[\mathbf{Z}] = [\mathbf{A}] - [V]\,[\mathbf{D}]\,, \tag{D.3}$$

$$[\mathbf{D}] = ([V]^T[M]\,[\mathbf{A}] + [\mathbf{A}]^T[M]\,[V])/2\,, \tag{D.4}$$

$$[\mathbf{A}] = [\mathbf{U}] - [B]\,\{[B]^T\,[\mathbf{U}]\}\,, \tag{D.5}$$

$$[\mathbf{U}] = [\mathbf{G}(\beta)]\,[\sigma]\,. \tag{D.6}$$

in which the random matrix $[\mathbf{G}(\beta)]$ is constructed by using the stochastic model presented in this Appendix.

D.2. Introduction of a continuous ROB

In this section, we introduce a continuous ROB corresponding to the computational ROB $[V] \in \mathcal{S}_{N,n} \subset \mathbb{S}_{N,n}$ such that $[V]^T[M]\,[V] = [I_n]$ and $[B]^T[V] = [0_{N_{\mathrm{CD}},n}]$, in which $n \geq 1$ is the dimension of the ROB, $[M]$ is a given matrix in $\mathbb{M}_N^+$, and $[B]$ is a given matrix in $\mathbb{M}_{N,N_{\mathrm{CD}}}$ such that $[B]^T\,[B] = [I_{N_{\mathrm{CD}}}]$. In order to simplify the presentation, we consider a continuous ROB made up of $\mathbb{R}^m$-valued fields and constraint equations defined by zero Dirichlet conditions. The extension to general constraint equations is straightforward.

(i) *Notations*. Let $d \geq 1$ be an integer. Let $\mathbf{x} = (x_1, \ldots, x_d)$ be the generating point of an open bounded domain $\Omega \subset \mathbb{R}^d$ and let $d\mathbf{x} = dx_1 \ldots dx_d$ be the Lebesgue measure. Let $\mathbf{x} \mapsto \rho(\mathbf{x})$ be a given bounded function on Ω with valued in $]0\,,+\infty[$. Let $m \geq 1$ be an integer and let $\mathrm{L}^2 := L^2_{\rho\,d\mathbf{x}}(\Omega, \mathbb{R}^m)$ be the space of all the $\mathbb{R}^m$-valued functions $\mathbf{x} \mapsto \mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \ldots, v_m(\mathbf{x}))$ that are square integrable on Ω with respect to the measure $\rho(\mathbf{x})\,d\mathbf{x}$, equipped with the usual inner product and the associated norm,

$$<\mathbf{u}\,,\mathbf{v}>_{\mathrm{L}^2} = \int_\Omega \mathbf{u}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})\,\rho(\mathbf{x})\,d\mathbf{x} \quad , \quad \|\mathbf{v}\|_{\mathrm{L}^2} = \{<\mathbf{v}\,,\mathbf{v}>_{L^2}\}^{1/2}\,, \tag{D.7}$$

in which $\mathbf{u}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}) = \sum_{\ell=1}^{m} u_\ell(\mathbf{x})\, v_\ell(\mathbf{x})$.

(ii) *Continuous ROB*. The continuous ROB is represented by the field $\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})] = [\mathbf{v}^1(\mathbf{x}) \ldots \mathbf{v}^n(\mathbf{x})]$ defined on Ω with values in $\mathbb{M}_{m,n}$ (for fixed $\mathbf{x}$, the n columns of $[\mathcal{V}(\mathbf{x})]$ are the n vectors $\mathbf{v}^1(\mathbf{x}), \ldots, \mathbf{v}^n(\mathbf{x})$). It is assumed that, for all k in $\{1, \ldots, n\}$, function $\mathbf{v}^k$ belongs to a subset $\mathcal{C}_{\text{ad}} \subset \mathrm{L}^2$, which is such that the trace of $\mathbf{v}^k$ on the boundary $\partial\Omega$ of Ω exits (often, $\mathcal{C}_{\text{ad}}$ is a Sobolev space such as $\mathrm{H}^p_{\rho\, d\mathbf{x}}(\Omega, \mathbb{R}^m)$ with $p \geq 1$) . Let $[\mathcal{C}_{\text{ad}}]$ be the corresponding subset of functions $\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})]$ defined on Ω with values in $\mathbb{M}_{m,n}$, such that $\mathbf{v}^k$ belongs to $\mathcal{C}_{\text{ad}}$ for all k,

$$[\mathcal{C}_{\text{ad}}] \;=\; \{\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})] = [\mathbf{v}^1(\mathbf{x}) \ldots \mathbf{v}^n(\mathbf{x})] : \Omega \to \mathbb{M}_{m,n} \;,\; \mathbf{v}^k \in \mathcal{C}_{\text{ad}} \subset \mathrm{L}^2 \,,\, \forall k = 1, \ldots, n\} \,. \quad \text{(D.8)}$$

Therefore, if $[\mathcal{V}]$ belongs to $[\mathcal{C}_{\text{ad}}]$, then the trace $[\mathcal{V}]_{|\partial\Omega}$ exists. In addition, it is assumed that $\{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$ constitutes an orthonormal family of functions in L^2, that is to say, $< \mathbf{v}^k \,, \mathbf{v}^{k'} >_{\mathrm{L}^2} = \delta_{kk'}$. Consequently, we have

$$\int_\Omega [\mathcal{V}(\mathbf{x})]^T [\mathcal{V}(\mathbf{x})]\, \rho(\mathbf{x})\, d\mathbf{x} \;=\; [I_n] \,. \quad \text{(D.9)}$$

The manifold $\mathbb{S}^\infty_{m,n}$ is defined as the set of all the functions $\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})]$ belonging to $[\mathcal{C}_{\text{ad}}]$ and verifying Eq. (D.9), that is to say,

$$\mathbb{S}^\infty_{m,n} \;=\; \{\, [\mathcal{V}] \in [\mathcal{C}_{\text{ad}}] \;\;,\;\; \int_\Omega [\mathcal{V}(\mathbf{x})]^T [\mathcal{V}(\mathbf{x})]\, \rho(\mathbf{x})\, d\mathbf{x} = \; [I_n] \,\} \,. \quad \text{(D.10)}$$

(iii) *Constraint equations*. As explained before, we consider the simple case for which a zero Dirichlet condition is given on a part Γ_0 of the boundary $\partial\Omega$ of Ω, that is to say, $\mathbf{v}^k(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$ in Γ_0 and for all $k = 1, \ldots, n$. The constraint equations can then be rewritten, for $[\mathcal{V}] \in [\mathcal{C}_{\text{ad}}]$, as

$$[\mathcal{V}(\mathbf{x})] = [0_{m,n}] \quad , \quad \forall\, \mathbf{x} \in \Gamma_0 \,. \quad \text{(D.11)}$$

(iv) *Definition of the manifold $\mathcal{S}^\infty_{m,n} \subset \mathbb{S}^\infty_{m,n}$ for the continuous ROB*. The manifold $\mathcal{S}^\infty_{m,n}$ is defined as the subset of $\mathbb{S}^\infty_{m,n}$ such that Eq. (D.11) holds:

$$\mathcal{S}^\infty_{m,n} = \{\, [\mathcal{V}] \in \mathbb{S}^\infty_{m,n} \;\;,\;\; [\mathcal{V}(\mathbf{x})] = [0_{m,n}] \,,\, \forall\, \mathbf{x} \in \Gamma_0 \,\} \;\subset\; \mathbb{S}^\infty_{m,n} \,. \quad \text{(D.12)}$$

(v) *Tangent vector space $T_{\mathcal{V}}\mathbb{S}^\infty_{m,n}$ to $\mathbb{S}^\infty_{m,n}$ at the point $[\mathcal{V}]$ in $\mathbb{S}^\infty_{m,n}$*. Such a tangent vector space is defined as the set of all the functions $[\mathcal{Z}] = \{\mathbf{x} \mapsto [\mathcal{Z}(\mathbf{x})]\}$ belonging to $[\mathcal{C}_{\text{ad}}]$ such that

$$\int_\Omega ([\mathcal{V}(\mathbf{x})]^T\, [\mathcal{Z}(\mathbf{x})] + [\mathcal{Z}(\mathbf{x})]^T\, [\mathcal{V}(\mathbf{x})])\, \rho(\mathbf{x})\, d\mathbf{x} = [0_{n,n}] \,. \quad \text{(D.13)}$$

D.3. Construction of the stochastic model of a continuous SROB

(i) *Continuous SROB*. For a given continuous ROB that is represented by the deterministic field $[\mathcal{V}] = \{\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})]\}$ belonging to manifold $\mathcal{S}^\infty_{m,n}$, the associated continuous SROB consists in introducing a random field $[\boldsymbol{\mathcal{W}}] = \{[\boldsymbol{\mathcal{W}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ that is a random variable defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$ with values in $\mathcal{S}^\infty_{m,n}$. Consequently, $[\boldsymbol{\mathcal{W}}]$ is a random field belonging to $[\mathcal{C}_{\text{ad}}]$ almost surely (a.s), which admits a trace on Γ_0 a.s, and which is such that

$$\int_\Omega [\boldsymbol{\mathcal{W}}(\mathbf{x})]^T [\boldsymbol{\mathcal{W}}(\mathbf{x})]\, \rho(\mathbf{x})\, d\mathbf{x} \;=\; [I_n] \quad a.s \quad , \quad [\boldsymbol{\mathcal{W}}(\mathbf{x})] = [0_{1,n}] \;,\; \forall\, \mathbf{x} \in \Gamma_0 \quad a.s \,. \quad \text{(D.14)}$$

The stochastic representation of random field $[\boldsymbol{\mathcal{W}}]$ is then defined by

$$[\boldsymbol{\mathcal{W}}(\mathbf{x})] = ([\mathcal{V}(\mathbf{x})] + s\, [\boldsymbol{\mathcal{Z}}(\mathbf{x})])\, [\boldsymbol{\mathcal{H}}_s(\boldsymbol{\mathcal{Z}})] \quad , \quad \forall\, \mathbf{x} \in \Omega \,, \quad \text{(D.15)}$$

$$[\boldsymbol{\mathcal{H}}_s(\boldsymbol{\mathcal{Z}})] = ([I_n] + s^2 \int_\Omega [\boldsymbol{\mathcal{Z}}(\mathbf{x})]^T\, [\boldsymbol{\mathcal{Z}}(\mathbf{x})]\, \rho(\mathbf{x})\, d\mathbf{x})^{-1/2} \,, \quad \text{(D.16)}$$

in which $[\boldsymbol{\mathcal{Z}}] = \{[\boldsymbol{\mathcal{Z}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is a random field defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, with values in $\mathbb{M}_{m,n}$, which belongs to tangent vector space $T_{\mathcal{V}}\mathbb{S}^{\infty}_{m,n}$ a.s, and which is such that

$$[\boldsymbol{\mathcal{Z}}(\mathbf{x})] = [0_{m,n}] \quad , \quad \forall \mathbf{x} \in \Gamma_0 \quad a.s\,. \tag{D.17}$$

The random matrix $[\boldsymbol{\mathcal{H}}_s(\boldsymbol{\mathcal{Z}})]$ is with values in $\mathbb{M}^+_n$. Using Eq. (D.15) with Eqs. (D.9), (D.13), and (D.16) yield the first equation in Eq. (D.14). Taking into account Eqs. (D.11) and (D.17), Eq. (D.15) yields the second equation in Eq. (D.14). Random field $\{[\boldsymbol{\mathcal{Z}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ can be written as

$$[\boldsymbol{\mathcal{Z}}(\mathbf{x})] = [\boldsymbol{\mathcal{A}}(\mathbf{x})] - [\mathcal{V}(\mathbf{x})]\,[\boldsymbol{\mathcal{D}}(\boldsymbol{\mathcal{A}})] \quad , \quad \forall \mathbf{x} \in \Omega\,, \tag{D.18}$$

$$[\boldsymbol{\mathcal{D}}(\boldsymbol{\mathcal{A}})] = \int_{\Omega} \frac{1}{2}([\mathcal{V}(\mathbf{x})]^T\,[\boldsymbol{\mathcal{A}}(\mathbf{x})] + [\boldsymbol{\mathcal{A}}(\mathbf{x})]^T\,[\mathcal{V}(\mathbf{x})])\,\rho(\mathbf{x})\,d\mathbf{x}\,, \tag{D.19}$$

in which $[\boldsymbol{\mathcal{D}}(\boldsymbol{\mathcal{A}})]$ is a random matrix with values in $\mathbb{M}^S_n$ and where $[\boldsymbol{\mathcal{A}}] = \{[\boldsymbol{\mathcal{A}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ is a random field defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, whose trajectories belong to $[\mathcal{C}_{\text{ad}}]$ a.s, and are such that

$$[\boldsymbol{\mathcal{A}}(\mathbf{x})] = [0_{m,n}] \quad , \quad \forall \mathbf{x} \in \Gamma_0 \quad a.s\,. \tag{D.20}$$

It should be noted that random field $[\boldsymbol{\mathcal{A}}]$ does not belong to tangent vector space $T_{\mathcal{V}}\mathbb{S}^{\infty}_{m,n}$. Substituting Eq. (D.18) with (D.19) into the left-hand side of Eq. (D.13) yields that Eq. (D.13) is satisfied and consequently, $[\boldsymbol{\mathcal{Z}}]$ belongs to tangent vector space $T_{\mathcal{V}}\mathbb{S}^{\infty}_{m,n}$ a.s.

(ii) *Construction of random field* $[\boldsymbol{\mathcal{A}}]$ *for obtaining the required regularity of the trajectories and the Dirichlet conditions*. The regularity of the trajectories of random field $[\boldsymbol{\mathcal{A}}]$ is defined by the regularity of functions in $[\mathcal{C}_{\text{ad}}]$. In addition, the trajectories must verify the zero Dirichlet condition given by Eq. (D.20). In order to construct a general theory, we propose the following approach. Let $\overline{\Omega} = \Omega \cup \partial\Omega$ be the closure of open bounded domain Ω for which its boundary $\partial\Omega = \Gamma_0 \cup \Gamma_1$ with $\Gamma_0 \cap \Gamma_1 = \emptyset$ is assumed to be smooth. Let be $\Omega_0 = \overline{\Omega} \backslash \Gamma_0$. Let $\mathbf{x} \mapsto \mathbb{1}_{\Omega_0}(\mathbf{x})$ be the function defined on $\overline{\Omega} = \Omega_0 \cup \Gamma_0$ such that $\mathbb{1}_{\Omega_0}(\mathbf{x}) = 1$ for all $\mathbf{x}$ in Ω_0 and $\mathbb{1}_{\Omega_0}(\mathbf{x}) = 0$ for all $\mathbf{x}$ in Γ_0. Note that function $\mathbb{1}_{\Omega_0}$ is not continuous in $\overline{\Omega}$. We propose to construct the random field $[\boldsymbol{\mathcal{A}}]$ as

$$[\boldsymbol{\mathcal{A}}(\mathbf{x})] = \mathbb{1}^{\varepsilon_r}_{\Omega_0}(\mathbf{x})\,[\boldsymbol{\mathcal{U}}(\mathbf{x})] \quad , \quad \forall \mathbf{x} \in \overline{\Omega}\,, \tag{D.21}$$

in which

- $\mathbf{x} \mapsto \mathbb{1}^{\varepsilon_r}_{\Omega_0}(\mathbf{x})$ is a real-valued function defined on $\overline{\Omega}$, which corresponds to a regularization of function $\mathbf{x} \mapsto \mathbb{1}_{\Omega_0}(\mathbf{x})$, and where $\varepsilon_r > 0$ is a small parameter that allows for driving the regularization. Let $\mathcal{V}^0_{\varepsilon_r}$ be a neighborhood of Γ_0, controlled by ε_r, which is defined as a subset of $\overline{\Omega}$, such that $\Gamma_0 \subset \mathcal{V}^0_{\varepsilon_r}$, and detailed in the Remark given after. The regularization is constructed in order that $\mathbb{1}^{\varepsilon_r}_{\Omega_0}(\mathbf{x})$ is equal (in the meaning of the regularization) to 0 if $\mathbf{x} \in \Gamma_0$, is equal to 1 if $\mathbf{x} \in \overline{\Omega} \backslash \mathcal{V}^0_{\varepsilon_r}$, is such that function $\mathbb{1}^{\varepsilon_r}_{\Omega_0}$ is indefinitely continuously differentiable on $\overline{\Omega}$, and is such that, if $\varepsilon_r \to 0$, then $\mathbb{1}^{\varepsilon_r}_{\Omega_0} \to \mathbb{1}_{\Omega_0}$ in L^1.

- $[\boldsymbol{\mathcal{U}}] = \{[\boldsymbol{\mathcal{U}}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ is a random field defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by $\mathbb{R}^d$, with values in $\mathbb{M}_{m,n}$, which will be constructed after in order that its trajectories are indefinitely continuously differentiable functions on $\mathbb{R}^d$ a.s. The random field $\{\mathbb{1}_{\Omega_0}(\mathbf{x})\,[\boldsymbol{\mathcal{U}}(\mathbf{x})], \mathbf{x} \in \Omega\}$ belongs to $[\mathcal{C}_{\text{ad}}]$ but the trajectories are discontinuous on Γ_0. This is the reason why the regularization $\mathbb{1}^{\varepsilon_r}_{\Omega_0}$ of $\mathbb{1}_{\Omega_0}$ has been introduced (due to the presence of a the zero Dirichlet condition on Γ_0). For ε_r sufficiently small, the trajectories of $\{\mathbb{1}^{\varepsilon_r}_{\Omega_0}(\mathbf{x})\,[\boldsymbol{\mathcal{U}}(\mathbf{x})], \mathbf{x} \in \overline{\Omega}\}$ are indefinitely continuously differentiable on $\overline{\Omega}$ and are such that $[\boldsymbol{\mathcal{A}}(\mathbf{x})] \simeq [0_{m,n}]$ for all $\mathbf{x} \in \Gamma_0$. Consequently, for a sufficiently small value of $\varepsilon_r > 0$, random field $[\boldsymbol{\mathcal{A}}]$ defined by Eq. (D.21), verifies the zero Dirichlet condition on Γ_0 (defined by Eq. (D.20)) and its trajectories (that are indefinitely continuous differentiable on $\overline{\Omega}$) thus belongs to $[\mathcal{C}_{\text{ad}}]$.

Remark about the construction of the regularization. This problem is relatively difficult and there are several possible methods. Hereinafter, we propose an approach for which the numerical cost is low.

- Let $\mathbf{x} \mapsto \mathbb{1}_{\Omega}^{\pm}(\mathbf{x})$ be the function defined almost everywhere on $\mathbb{R}^d$ such that $\mathbb{1}_{\Omega}^{\pm}(\mathbf{x}) = 1$ for all $\mathbf{x}$ in Ω and $\mathbb{1}_{\Omega}^{\pm}(\mathbf{x}) = -1$ for all $\mathbf{x}$ in $\mathbb{R}^d \backslash \Omega$. The first step consists in introducing a usual regularization $\mathbb{1}_{\Omega}^{\varepsilon_r}$ of function $\mathbb{1}_{\Omega}^{\pm}$, which can be constructed, for all $\mathbf{x}$ in $\mathbb{R}^d$, as

$$\mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x}) = (p_{\varepsilon_r} \star \mathbb{1}_{\Omega}^{\pm})(\mathbf{x}) = \int_{\mathbb{R}^d} p_{\varepsilon_r}(\mathbf{y})\, \mathbb{1}_{\Omega}^{\pm}(\mathbf{x}-\mathbf{y})\, d\mathbf{y} = E\{\mathbb{1}_{\Omega}^{\pm}(\mathbf{x} - \mathbf{Y}^{\varepsilon_r})\}\,, \tag{D.22}$$

  in which $\mathbf{Y}^{\varepsilon_r} = (Y_1^{\varepsilon_r}, \ldots, Y_d^{\varepsilon_r})$ is the Gaussian $\mathbb{R}^d$-valued random variable whose probability density function with respect to $d\mathbf{y}$ is written as

$$p_{\varepsilon_r}(\mathbf{y}) = \frac{1}{\varepsilon_r^d (2\pi)^{d/2}}\, \exp(-\frac{\|\mathbf{y}\|^2}{2{\varepsilon_r}^2})\,. \tag{D.23}$$

  Therefore, the real-valued random variables $Y_1^{\varepsilon_r}, \ldots, Y_d^{\varepsilon_r}$ are independent, Gaussian, centered, and with variance ε_r^2. Taking into account Eq. (D.22), it should be noted that, for all fixed $\mathbf{x}$ in $\mathbb{R}^d$, $\mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x})$ can be estimated by the Monte Carlo method.

- For ε_r sufficiently small, let $\mathcal{V}_{\varepsilon_r}^{\pm}$ be the open neighborhood of $\partial\Omega$ in $\mathbb{R}^d$ associated with the regularization such that (1) $\mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x}) \sim 0$ if $\mathbf{x} \in \partial\Omega$, (2) $\mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x}) \sim 1$ if $\mathbf{x} \in \overline{\Omega} \backslash \mathcal{V}_{\varepsilon_r}^{+}$ in which $\mathcal{V}_{\varepsilon_r}^{+} = \overline{\Omega} \cap \mathcal{V}_{\varepsilon_r}^{\pm}$, and (3) $\mathbf{x} \mapsto \mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x})$ is indefinitely continuously differentiable on $\overline{\Omega}$. Note that $\partial\Omega \subset \mathcal{V}_{\varepsilon_r}^{+}$. Let $\mathcal{V}_{\varepsilon_r}^{0}$ be the subset of $\mathcal{V}_{\varepsilon_r}^{+}$ such that $\Gamma_0 \subset \mathcal{V}_{\varepsilon_r}^{0}$ and which is defined as a neighborhood of Γ_0.

- The regularization function $\mathbf{x} \mapsto \mathbb{1}_{\Omega_0}^{\varepsilon_r}(\mathbf{x})$ on $\overline{\Omega}$ of function $\mathbf{x} \mapsto \mathbb{1}_{\Omega_0}(\mathbf{x})$ on $\overline{\Omega}$ is constructed as follows,

$$\mathbb{1}_{\Omega_0}^{\varepsilon_r}(\mathbf{x}) \sim 1\ ,\ \forall\, \mathbf{x} \in \overline{\Omega} \backslash \mathcal{V}_{\varepsilon_r}^{0} \quad , \quad \mathbb{1}_{\Omega_0}^{\varepsilon_r}(\mathbf{x}) = \mathbb{1}_{\Omega}^{\varepsilon_r}(\mathbf{x})\ ,\ \forall\, \mathbf{x} \in \mathcal{V}_{\varepsilon_r}^{0}\,. \tag{D.24}$$

  Consequently, $\mathbb{1}_{\Omega_0}^{\varepsilon_r}(\mathbf{x}) \sim 0$ if $\mathbf{x} \in \Gamma_0$.

(iii) *Construction of random field* $[\boldsymbol{\mathcal{U}}]$. Random field $[\boldsymbol{\mathcal{U}}] = [\boldsymbol{\mathcal{U}}^1 \ldots \boldsymbol{\mathcal{U}}^n]$ indexed by $\mathbb{R}^d$ values in $\mathbb{M}_{m,n}$ is constructed as a second-order, centered, mean-square homogeneous (stationary), and mean-square continuous random field for which the fourth-order tensor-valued autocorrelation function $(\mathbf{x}, \mathbf{x}') \mapsto \{R_{\boldsymbol{\mathcal{U}}}(\mathbf{x}, \mathbf{x}')\}_{\ell k \ell' k'} = E\{[\boldsymbol{\mathcal{U}}(\mathbf{x})]_{\ell k} [\boldsymbol{\mathcal{U}}(\mathbf{x}')]_{\ell' k'}\}$ that is defined on $\mathbb{R}^d \times \mathbb{R}^d$, is written, for ℓ and ℓ' in $\{1, \ldots, m\}$ and for k and k' in $\{1, \ldots, n\}$, as

$$\{R_{\boldsymbol{\mathcal{U}}}(\mathbf{x}, \mathbf{x}')\}_{\ell k \ell' k'} = R(\mathbf{x} - \mathbf{x}'; \beta)\, \delta_{\ell \ell'}\, [c_n]_{k k'} \quad , \quad (\mathbf{x}, \mathbf{x}') \in \mathbb{R}^d \times \mathbb{R}^d\,, \tag{D.25}$$

in which $[c_n] = [\sigma]^T [\sigma] \in \mathbb{M}_n^+$ with $[\sigma] \in \mathbb{M}_n^u$ the upper triangular matrix with positive diagonal entries (see Eq. (3.16)). The real function $\boldsymbol{\eta} \mapsto R(\boldsymbol{\eta}; \beta)$ is an autocorrelation-type function which is constructed to be indefinitely continuously differentiable on $\mathbb{R}^d$ and to go to zero as $\|\boldsymbol{\eta}\|$ goes to $+\infty$. Such a condition will be fulfilled thanks to the construction of the stochastic model of the random field $\{\mathcal{Q}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ presented hereinafter (see Paragraph (iv)).The choice defined by Eq. (D.25) for the autocorrelation function of random field $[\boldsymbol{\mathcal{U}}]$ has voluntary been chosen as simple in order to be in capability to use the SROM for very large values of N. This algebraic representation for $R_{\boldsymbol{\mathcal{U}}}(\mathbf{x}, \mathbf{x}')$ implies the following properties:

- Matrix $[c_n]$ allows for controlling the correlation between the n random fields $\boldsymbol{\mathcal{U}}^1, \ldots \boldsymbol{\mathcal{U}}^n$ with values in $\mathbb{R}^m$, using $n(n+1)/2$ hyperparameters (upper triangular matrix $[\sigma]$).
- The m components $\{\mathcal{U}_1^k(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}, \ldots, \{\mathcal{U}_m^k(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ of $\mathbb{R}^m$-valued random field $\{\boldsymbol{\mathcal{U}}^k(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ are not correlated. Such a choice is mainly done in order to not increase the number of hyperparameters in the stochastic modeling.

- The real-valued function $\boldsymbol{\eta} \mapsto R(\boldsymbol{\eta};\beta)$ allows for controlling the real-valued autocorrelation function of each $\mathbb{R}^m$-valued random field $\{\boldsymbol{\mathcal{U}}^k(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$, using only one real-valued hyperparameter β that will be related to the spatial correlation length of the field. With the proposed construction, this autocorrelation function is the same for each k.
- There is a version of random field $[\boldsymbol{\mathcal{U}}]$ whose trajectories are indefinitely continuously differentiable functions on $\mathbb{R}^d$ a.s. Consequently, there is a version $\{[\boldsymbol{\mathcal{W}}(\mathbf{x})], \mathbf{x} \in \overline{\Omega}\}$ of the SROB for which its trajectories belong to $[\mathcal{C}_{\text{ad}}]$ and verify the zero Dirichlet condition on Γ_0, and therefore, belong to manifold $\mathcal{S}^{\infty}_{m,n}$.

(iv) *Normalization of random field* $[\boldsymbol{\mathcal{U}}]$. Taking into account Eq. (D.25), random field $[\boldsymbol{\mathcal{U}}]$ can be rewritten as

$$[\boldsymbol{\mathcal{U}}(\mathbf{x})] = [\boldsymbol{\mathcal{G}}(\mathbf{x})]\,[\sigma] \quad , \quad \forall\, \mathbf{x} \in \mathbb{R}^d\,, \tag{D.26}$$

in which $\{[\boldsymbol{\mathcal{G}}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ is a random field defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by $\mathbb{R}^d$, with values in $\mathbb{M}_{m,n}$, which is a second-order, centered, and mean-square homogeneous random field such that its fourth-order tensor-valued autocorrelation function $(\mathbf{x}, \mathbf{x}') \mapsto \{R_{\boldsymbol{\mathcal{G}}}(\mathbf{x}, \mathbf{x}')\}_{\ell k \ell' k'} = E\{[\boldsymbol{\mathcal{G}}(\mathbf{x})]_{\ell k}[\boldsymbol{\mathcal{G}}(\mathbf{x}')]_{\ell' k'}\}$ defined on $\mathbb{R}^d \times \mathbb{R}^d$, is written, for ℓ and ℓ' in $\{1, \ldots, m\}$ and for k and k' in $\{1, \ldots, n\}$, as

$$\{R_{\boldsymbol{\mathcal{G}}}(\mathbf{x}, \mathbf{x}')\}_{\ell k \ell' k'} = R(\mathbf{x} - \mathbf{x}'; \beta)\, \delta_{\ell \ell'}\, \delta_{k k'} \quad , \quad (\mathbf{x}, \mathbf{x}') \in \mathbb{R}^d \times \mathbb{R}^d\,. \tag{D.27}$$

The stochastic model of random field $\{[\boldsymbol{\mathcal{G}}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ is constructed in order that:

- For all ℓ in $\{1, \ldots, m\}$ and for all k in $\{1, \ldots, n\}$, the random fields $\{[\boldsymbol{\mathcal{G}}(\mathbf{x})]_{\ell k}, \mathbf{x} \in \mathbb{R}^d\}$ are *independent copies* of a random field $\{\mathcal{Q}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by $\mathbb{R}^d$, with values in $\mathbb{R}$.
- The random field $\{\mathcal{Q}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ (that is defined hereinafter in Paragraph (v)) is second-order, centered, mean-square homogeneous, mean-square continuous, and there there is a version of random field $\mathcal{Q}$ whose trajectories are indefinitely continuously differentiable functions on $\mathbb{R}^d$ a.s.

Taking into account Eq. (D.27), it can be deduced that the real-valued autocorrelation function $(\mathbf{x}, \mathbf{x}') \mapsto R_{\mathcal{Q}}(\mathbf{x}, \mathbf{x}') = E\{\mathcal{Q}(\mathbf{x})\, \mathcal{Q}(\mathbf{x}')\}$ defined on $\mathbb{R}^d \times \mathbb{R}^d$, is written,

$$[R_{\mathcal{Q}}(\mathbf{x}, \mathbf{x}')] = R(\mathbf{x} - \mathbf{x}'; \beta) \quad , \quad (\mathbf{x}, \mathbf{x}') \in \mathbb{R}^d \times \mathbb{R}^d\,. \tag{D.28}$$

(v) *Stochastic modeling of random field* $\mathcal{Q}$. It is recalled that the objective is a construction of a SROB in the context of the nonparametric probabilistic approach of modeling errors and not the construction of a parametric probabilistic model of uncertain physical parameters of the HDM. This means that we have the freedom to choose an adapted stochastic model.

(v-1) *Explicit construction of autocorrelation function* $R(\boldsymbol{\eta};\beta)$ *of field* $\mathcal{Q}$. For all i in $\{1, \ldots, d\}$, let $L_i > 0$ be the maximum length of domain $\overline{\Omega}$ in direction i, defined by

$$L_i = \max_{(\mathbf{x}, \mathbf{x}') \in \overline{\Omega} \times \overline{\Omega}} |x_i - x'_i|\,. \tag{D.29}$$

The algebraic model of autocorrelation function $\boldsymbol{\eta} \mapsto R(\boldsymbol{\eta};\beta)$ defined on $\mathbb{R}^d$, with values in $\mathbb{R}$, and depending on hyperparameter β, is chosen as

$$R(\boldsymbol{\eta};\beta) = \zeta \times r_1(\eta_1;\beta) \times \ldots \times r_d(\eta_d;\beta) \quad , \quad \boldsymbol{\eta} = (\eta_1, \ldots, \eta_d)\,, \tag{D.30}$$

in which $\zeta > 0$ is a constant of normalization (that will be defined after), where for all $i = 1, \ldots, d$,

$$r_i(0;\beta) = 1 \quad ; \quad r_i(\eta_i;\beta) = \frac{4\mathcal{L}_i^2}{\pi^2 \eta_i^2}\, \sin^2\!\left(\frac{\pi \eta_i}{2\mathcal{L}_i}\right) \quad \text{for} \quad \eta_i \neq 0\,, \tag{D.31}$$

and where $\mathcal{L}_1, \ldots, \mathcal{L}_d$ are positive real numbers depending on η, which are defined by

$$\mathcal{L}_i = \beta\, L_i \quad , \quad i = 1, \ldots, d \quad , \quad \beta > 0\,. \tag{D.32}$$

Therefore, there exists a power spectral density function $\mathbf{k} = (k_1, \ldots, k_d) \mapsto S(\mathbf{k}; \beta)$, integrable from $\mathbb{R}^d$ into $\mathbb{R}^+$, which is written as

$$S(\mathbf{k}; \beta) = s_1(k_1; \beta) \times \ldots \times s_d(k_d; \beta)\,, \tag{D.33}$$

in which, for all $i = 1, \ldots, d$, the function $k_i \mapsto s_i(k_i; \beta)$ from $\mathbb{R}$ into $\mathbb{R}^+$ is written as

$$s_i(k_i; \beta) = \frac{\mathcal{L}_i}{\pi}\, \chi(k_i \frac{\mathcal{L}_i}{\pi})\,. \tag{D.34}$$

In Eq. (D.34), the function $\kappa \mapsto \chi(\kappa)$ is continuous from $\mathbb{R}$ into $\mathbb{R}^+$, has a compact support $[-1\,, 1]$, and is such that

$$\chi(0) = 1 \quad ; \quad \chi(-\kappa) = \chi(\kappa) \quad ; \quad \chi(\kappa) = 1 - \kappa \quad \text{for} \quad \kappa \in [0\,, 1]\,. \tag{D.35}$$

Consequently, $\mathbf{k} \mapsto S(\mathbf{k}; \beta)$ is a function with a compact support

$$\text{supp}\, S(.\,; \beta) = [-K_1, K_1] \times \ldots \times [-K_d, K_d] \quad , \quad K_i = \frac{\pi}{\mathcal{L}_i} = \frac{\pi}{\beta\, L_i}\,. \tag{D.36}$$

The spatial correlation length $\mathcal{L}_i^{\mathcal{Q}}$ of random field $\mathcal{Q}$, relative to the coordinate x_i, is written as

$$\mathcal{L}_i^{\mathcal{Q}} = \int_0^{+\infty} r_i(\eta_i)\, d\eta_i = \frac{2\pi}{2}\, s_i(0; \beta) = \mathcal{L}_i = \beta\, L_i\,. \tag{D.37}$$

Consequently, the parameters $\mathcal{L}_1, \ldots, \mathcal{L}_d$ represent the spatial correlation lengths of random field $\mathcal{Q}$. Since $\mathcal{L}_i = \beta\, L_i$, hyperparameter β allows for controlling these spatial correlation lengths.

(v-2) *Regularity of the trajectories of random field $\mathcal{Q}$.* Using the Kolmogorov lemma, it can be deduced that the random field $\mathcal{Q}$ admits a version whose trajectories are indefinitely continuously differentiable functions on $\mathbb{R}^d$ a.s.

(v-3) *Choice of a representation of random field $\mathcal{Q}$ adapted to the numerical simulation for large values of N_o.* The first question that has to be analyzed is the choice of a representation for $\mathcal{Q}$, which is adapted to the problem that we have to solve under the specificities given hereinafter.

- A major constraint is related to the fact that we have to construct numerical simulations of random field $\mathcal{Q}$ for a set of points $\mathbf{x}^1, \ldots, \mathbf{x}^{N_o}$ belonging to $\overline{\Omega} \subset \mathbb{R}^d$ for which N_o can be very large, that is to say, we have to construct a generator of independent realizations of a random vector $\mathbb{Q} = (\mathcal{Q}(\mathbf{x}^1), \ldots, \mathcal{Q}(\mathbf{x}^{N_o}))$ with values in $\mathbb{R}^{N_o}$.

- The construction of a generator of independent realizations of random field $\mathcal{Q}$ requires the construction of it probability law (which can be defined, for instance, by introducing the system of marginal probability distributions of $\mathcal{Q}$). The choice of the probability law for $\mathcal{Q}$ is not crucial, because we are interested in generating randomness for which only the spatial correlation length (represented by hyperparameter β) has to be controlled. Consequently, we are interested in constructing a representation of random field $\mathcal{Q}$, which is very efficient for the numerical simulations in high dimension (large value of N_o), without imposing a given probability law for random field $\mathcal{Q}$. Nevertheless, $\mathcal{Q}$ must be a second-order, mean-square homogeneous (mean-square stationary), centered random field for which its autocorrelation function is defined by Eqs. (D.30) to (D.32).

- In this framework for which there is no specification concerning the probability law, it could be assumed that $\mathcal{Q}$ is a Gaussian random field in order to construct a representation of $\mathcal{Q}$ (or more precisely, a representation for its discretized form $\mathbb{Q}$). A possible candidate for such a representation of $\mathbb{Q}$ would be $\mathbb{Q} = [L_{\mathbb{Q}}]\, \mathbb{G}$ in which the lower triangular $(N_o \times N_o)$ real matrix $[L_{\mathbb{Q}}]$ would come from the Cholesky factorization, $[C_{\mathbb{Q}}] = [L_{\mathbb{Q}}]\, [L_{\mathbb{Q}}]^T$, of the

covariance matrix $[C_{\mathbb{Q}}] \in \mathbb{M}^+_{N_o}$ of the centered random vector $\mathbb{Q}$. Matrix $[C_{\mathbb{Q}}]$ is such that $[C_{\mathbb{Q}}]_{j_o j'_o} = R(\mathbf{x}^{j_o} - \mathbf{x}^{j'_o})$, and $\mathbb{G}$ would be the second-order centered Gaussian vector with values in $\mathbb{R}^{N_o}$, for which the covariance matrix is $[I_{N_o}]$. For a large value of N_o, such a representation of $\mathbb{Q}$ could not be used because $[C_{\mathbb{Q}}]$ is a full big matrix. It should be noted that, from a numerical point of view, the construction of a Karhunen-Loeve expansion of random field $\mathcal{Q}$, would be equivalent to the principal component analysis of random vector $\mathbb{Q}$, which would consist in effectively constructing the big matrix $[C_{\mathbb{Q}}]$ and then, in extracting the dominant eigensubspace of $[C_{\mathbb{Q}}]$ using for instance the subspace iteration method or the Krylov iteration method. For large value of N_0, such a method could be very numerically expansive and would require a very large core memory.

- Always assuming that $\mathcal{Q}$ could be a Gaussian random field, another representation could be deduced from the use of the Shinozuka representation [68, 69] of random field $\mathcal{Q}$, which is a particular case of integral representations involving the stochastic spectral measure of random field $\mathcal{Q}$, and for which a detailed development with additional mathematical properties related to convergence of the representation can be found in [70]. Nevertheless, for $d \geq 2$ (in particular, for $d = 3$) such a representation would be very efficient if the multidimensional Fast Fourier Transform (FFT) could be used. In our case, since the set $\mathbf{x}^1, \ldots, \mathbf{x}^{N_o}$ of given points in $\overline{\Omega}$ does not constitute a structured mesh corresponding to a constant spatial sampling step in each direction, the multidimensional FFT cannot be used. However, as the autocorrelation function of $\mathcal{Q}$ is assumed to be separated (see Eq. (D.30)) and as the probability law of random field $\mathcal{Q}$ can be arbitrary (and consequently, can be non-Gaussian), we propose to generate $\mathcal{Q}$ in writing it as

$$\mathcal{Q}(\mathbf{x}) = \sqrt{\zeta} \times \mathcal{Q}_1(x_1) \times \ldots \times \mathcal{Q}_d(x_d) \quad , \quad \forall\, \mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d \,, \tag{D.38}$$

in which $\mathcal{Q}_1, \ldots, \mathcal{Q}_d$ are d independent Gaussian real-valued stochastic processes indexed by $\mathbb{R}$. This stochastic model allows for decreasing the numerical cost for the generation of independent realizations of random field $\mathcal{Q}$ at points $\mathbf{x}^1, \ldots, \mathbf{x}^{N_o}$ for very large large values of N_o. For i fixed in $\{1, \ldots, d\}$, $\mathcal{Q}_i$ is a Gaussian, second-order, mean-square stationary, centered stochastic process for which its autocorrelation function is

$$r_i(\eta_i; \beta) = E\{\mathcal{Q}_i(x_i + \eta_i)\, \mathcal{Q}_i(x_i)\} \quad , \quad \forall\, x_i \in \mathbb{R} \quad , \quad \forall\, \eta_i \in \mathbb{R} \,, \tag{D.39}$$

which is defined by Eqs. (D.31) and (D.32), and for which the correlation length is $\mathcal{L}_i = \beta\, L_i$. The power spectral density function of stochastic process $\mathcal{Q}_i$ is the function $k_i \mapsto s_i(k_i; \beta)$ defined by Eqs. (D.34), with compact support $[-K_i, K_i]$ with $K_i = \pi/\mathcal{L}_i = \pi/(\beta\, L_i)$, which is such that

$$r_i(\eta_i; \beta) = \int_{\mathbb{R}} e^{\mathrm{i}\, k_i \eta_i}\, s_i(k_i; \beta)\, dk_i \quad , \quad \forall\, \eta_i \in \mathbb{R} \,. \tag{D.40}$$

There exists a version of Gaussian stochastic process $\mathcal{Q}_i$ for which the trajectories are indefinitely continuously differentiable functions on $\mathbb{R}$ a.s. We can then use the Shinozuka representation for each Gaussian stochastic process $\mathcal{Q}_i$ indexed by $\mathbb{R}$. For the unstructured mesh that is considered, the algorithm complexity that was in N_o^d is now in N_o. For the numerical calculation, the series of the Shinozuka representation is truncated to a finite number ν_p of terms.

(v-4) *Representation of random field $\mathcal{Q}_i$.* Let $p \geq 1$ be an integer and let $\nu_p = 2\, p$. Let $\kappa_1, \ldots, \kappa_{\nu_p}$ be the sampling points of the interval $[-1\,, 1]$ with the constant step $\Delta = 2/\nu_p$ such that

$$\kappa_\nu = -1 + \left(\nu - \frac{1}{2}\right) \frac{2}{\nu_p} \quad , \quad \nu = 1, \ldots, \nu_p \,. \tag{D.41}$$

Let $\Sigma_1, \ldots, \Sigma_{\nu_p}$ be the positive real numbers defined by

$$\Sigma_\nu = \frac{2}{\nu_p}\, \chi(\kappa_\nu) \quad , \quad \nu = 1, \ldots, \nu_p \,, \tag{D.42}$$

in which $\kappa \mapsto \chi(\kappa)$ is defined by Eq. (D.35). We then have the following ν_p-order representation $\mathcal{Q}_i^{\nu_p}$ of $\mathcal{Q}_i$,

$$\mathcal{Q}_i^{\nu_p}(x_i) = \sum_{\nu=1}^{\nu_p} \sqrt{2\,\Sigma_\nu}\,\sqrt{-\log \Psi_\nu^i}\,\cos\{\Phi_\nu^i + \frac{\pi}{\mathcal{L}_i}\kappa_\nu\, x_i\} \quad , \quad \forall\, x_i \in \mathbb{R}\,, \tag{D.43}$$

in which $\{\Psi_1^i, \ldots, \Psi_{\nu_p}^i\}$ is a set of independent uniform random variables on $[0\,,1]$, where $\{\Phi_1^i, \ldots, \Phi_{\nu_p}^i\}$ is a set of independent uniform random variables on $[0\,, 2\pi]$, and where the $2\,d\,\nu_p$ random variables $\{\Psi_1^1, \ldots, \Psi_{\nu_p}^1\}, \ldots, \{\Psi_1^d, \ldots, \Psi_{\nu_p}^d\}$, $\{\Phi_1^1, \ldots, \Phi_{\nu_p}^1\}, \ldots, \{\Phi_1^d, \ldots, \Phi_{\nu_p}^d\}$ are independent. For any value of $\nu_p = 2\,p$ fixed, and for all i in $\{1, \ldots, d\}$, the stochastic process $\{\mathcal{Q}_i^{\nu_p}(x_i), x_i \in \mathbb{R}\}$ is Gaussian and if ν_p goes to $+\infty$, then the sequence $\{R_{\mathcal{Q}_i^{\nu_p}}\}_{\nu_p}$ of autocorrelation functions goes to autocorrelation function $r_i(.\,;\beta)$ of stochastic process $\{\mathcal{Q}_i(x_i), x_i \in \mathbb{R}\}$.

D.4. Finite element discretization

(i) *Notations for the finite element discretization*. Let $\mathbf{x}^1, \ldots, \mathbf{x}^{N_o}$ be the N_o nodes of the finite element mesh of domain $\overline{\Omega}$. The numbering of the nodes is assumed to be such that the first $N_{\text{CD}} < N_o$ nodes, $\mathbf{x}^1, \ldots, \mathbf{x}^{N_{\text{CD}}}$, correspond to all the nodes that belong to $\Gamma_0 \subset \partial\Omega$ on which the zero Dirichlet condition is written. Let $[\mathcal{N}(\mathbf{x})]$ be the row matrix in $\mathbb{M}_{m,N}$ of the assembled interpolation functions of all the finite elements. Let $N = m \times N_o$ be the number of degrees of freedom (before applying the zeros Dirichlet conditions). The finite element discretization of any continuous function $\mathbf{x} \mapsto \mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), \ldots, u_m(\mathbf{x}))$ from $\overline{\Omega}$ into $\mathbb{R}^m$ (which does not verify the Dirichlet condition on Γ_0), is the function $\mathbf{x} \mapsto \mathbf{u}^{(N)}(\mathbf{x})$ such that

$$\mathbf{u}^{(N)}(\mathbf{x}) = [\mathcal{N}(\mathbf{x})]\,\mathbb{U} \quad , \quad \forall\, \mathbf{x} \in \overline{\Omega}\,, \tag{D.44}$$

in which $\mathbb{U} = (\mathbb{u}_1, \ldots, \mathbb{u}_N)$ is the vector in $\mathbb{R}^N$ of the N DOFs $\mathbb{u}_j$ such that

$$\mathbb{u}_j = u_\ell(\mathbf{x}^{j_o}) \quad , \quad j \in \{1, \ldots, N\}\,, \tag{D.45}$$

in which the DOF $j = (\ell, j_o)$ is associated with the component ℓ of node j_o. Let us consider another continuous function $\mathbf{x} \mapsto \mathbf{a}(\mathbf{x}) = (a_1(\mathbf{x}), \ldots, a_m(\mathbf{x}))$ from $\overline{\Omega}$ into $\mathbb{R}^m$, which verifies the Dirichlet condition $\mathbf{a}(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$ in Γ_0. Let $\mathbf{a}^{(N)}(\mathbf{x}) = [\mathcal{N}(\mathbf{x})]\,\mathbb{A}$ be its finite element discretization in which $\mathbb{A} = (\mathbb{a}_1, \ldots, \mathbb{a}_N) \in \mathbb{R}^N$ with $\mathbb{a}_j = a_\ell(\mathbf{x}^{j_o})$. Consequently, we have

$$\mathbb{a}_j = 0 \quad , \quad \forall\, j \in \{1, \ldots, N_{\text{CD}}\}\,. \tag{D.46}$$

The parameterization of $\mathbb{A} \in \mathbb{R}^N$ that satisfies the Dirichlet conditions defined by Eq. (D.46), expressed as a function of any vector $\mathbb{U}$ in $\mathbb{R}^N$ can then be written as

$$\mathbb{A} = \mathbb{U} - [B]\,[B]^T\mathbb{U} \quad , \quad \forall\, \mathbb{U} \in \mathbb{R}^N\,, \tag{D.47}$$

in which $[B]$ is the matrix in $\mathbb{M}_{N,N_{\text{CD}}}$ such that $[B]_{ji} = \delta_{ji}$ for all $j \in \{1, \ldots, N\}$ and $i \in \{1, \ldots, N_{\text{CD}}\}$. It can be seen that

$$[B]^T[B] = [I_{N_{\text{CD}}}]\,. \tag{D.48}$$

(ii) *Finite element discretization of the continuous ROB*. Introducing the matrix $[V] = [\mathbf{v}^1 ... \mathbf{v}^n] \in \mathbb{M}_{N,n}$ with $\mathbf{v}^k = (v_1^k, \ldots, v_N^k) \in \mathbb{R}^N$, the finite element approximation, $\mathbf{x} \mapsto [\mathcal{V}^{(N)}(\mathbf{x})]$, of continuous ROB $\mathbf{x} \mapsto [\mathcal{V}(\mathbf{x})]$ is written as

$$[\mathcal{V}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[V] \quad , \quad [V] \in \mathbb{M}_{N,n} \quad , \quad \forall\, \mathbf{x} \in \overline{\Omega}\,. \tag{D.49}$$

We have now to verify that, if $[\mathcal{V}]$ belongs to $\mathcal{S}_{m,n}^{\infty}$, then $[V]$ belongs to $\mathcal{S}_{N,n}$. Substituting Eq. (D.49) into Eq. (D.9) yields

$$[V]^T[M]\,[V] = [I_n] \quad , \quad [M] = \int_\Omega [\mathcal{N}(\mathbf{x})]^T[\mathcal{N}(\mathbf{x})]\,\rho(\mathbf{x})\,d\mathbf{x} \in \mathbb{M}_n^+\,. \tag{D.50}$$

Equation (D.11) yields $[\mathcal{V}(\mathbf{x}^1)] = \ldots = [\mathcal{V}(\mathbf{x}^{N_{\text{CD}}})] = [0_{m,n}]$. Since $\mathbf{v}_j^k = v_\ell^k(\mathbf{x}^{j_o})$ with $j = (\ell, j_o)$ and as $[B]_{ji} = \delta_{ji}$ for all $j \in \{1, \ldots, N\}$ and $i \in \{1, \ldots, N_{\text{CD}}\}$, it can be deduced that

$$[B]^T [V] = [0_{N_{\text{CD}},n}] \,. \tag{D.51}$$

Equations (D.50) and (D.51) show that the computational ROB belongs to $\mathcal{S}_{N,n}$ (see Eq. (3.8)). The finite element approximation of a field $\mathbf{x} \mapsto [\mathcal{Z}(\mathbf{x})]$ continuous from $\overline{\Omega}$ into $\mathbb{M}_{m,n}$, which belongs to tangent vector space $T_{\mathcal{V}}\mathbb{S}^\infty_{m,n}$, is written as

$$[\mathcal{Z}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[Z] \quad , \quad [Z] \in \mathbb{M}_{N,n} \quad , \quad \forall\, \mathbf{x} \in \overline{\Omega} \,. \tag{D.52}$$

We have to verify that $[Z]$ belongs to tangent vector space $T_V\mathbb{S}_{N,n}$ to $\mathbb{S}_{N,n}$ at the point $[V]$ in $\mathbb{S}_{N,n}$. Substituting Eqs. (D.49) and (D.52) into Eq. (D.13) yields

$$[V]^T\,[M]\,[Z] + [Z]^T\,[M]\,[V] = [0_{n,n}] \,. \tag{D.53}$$

Equation (D.53) shows that $[Z]$ belongs to $T_V\mathbb{S}_{N,n}$ (see (3.2)).

(iii) *Finite element discretization of the continuous SROB.* Let $\{[\boldsymbol{\mathcal{F}}(\mathbf{x})], \mathbf{x} \in \overline{\Omega}\}$ be a random field representing one of the random fields $[\boldsymbol{\mathcal{W}}]$, $[\boldsymbol{\mathcal{Z}}]$, $[\boldsymbol{\mathcal{A}}]$, $[\boldsymbol{\mathcal{U}}]$, or $[\boldsymbol{\mathcal{G}}]$ defined in Section D.3. The finite element approximation of random field $[\boldsymbol{\mathcal{F}}]$ is written as

$$[\boldsymbol{\mathcal{F}}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[\mathbf{F}] \quad , \quad [\mathbf{F}] \in \mathbb{M}_{N,n} \quad , \quad \forall\, \mathbf{x} \in \overline{\Omega} \,, \tag{D.54}$$

in which $[\mathbf{F}]$ is the random matrix corresponding to the random matrices $[\mathbf{W}]$, $[\mathbf{Z}]$, $[\mathbf{A}]$, $[\mathbf{U}]$, or $[\mathbf{G}(\beta)]$ that are in Eqs. (D.1) to (D.6). The finite element method is such that, for all $j = 1, \ldots, N$ and for all $k = 1, \ldots, n$, we have $[\mathbf{F}]_{jk} = [\boldsymbol{\mathcal{F}}(\mathbf{x}^{j_o})]_{\ell k}$ with $j = (\ell, j_o)$. Substituting $[\boldsymbol{\mathcal{Z}}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[\mathbf{Z}]$ into Eq. (D.16) yields $[\boldsymbol{\mathcal{H}}_s(\boldsymbol{\mathcal{Z}}^{(N)})] = [H_s(\mathbf{Z})]$ in which $[H_s(\mathbf{Z})]$ is defined by Eq. (D.2). Similarly, substituting $[\boldsymbol{\mathcal{A}}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[\mathbf{A}]$ and $[\mathcal{V}^{(N)}(\mathbf{x})] = [\mathcal{N}(\mathbf{x})]\,[V]$ into Eq. (D.19) yields $[\boldsymbol{\mathcal{D}}(\boldsymbol{\mathcal{A}}^{(N)})] = [\mathbf{D}]$ in which $[\mathbf{D}]$ is defined by Eq. (D.4). Taking $\mathbf{x} = \mathbf{x}^{j_o}$ for $j_o = 1, \ldots, N_o$ into Eqs. (D.15), (D.18), and (D.26) yields Eqs. (D.1), (D.3), and (D.6). We have then proved that the finite element discretization of the continuous SROB introduced above correspond to the computational SROB constructed in Section 3.3. Nevertheless, we have to precise the construction of the stochastic model of random matrix $[\mathbf{G}(\beta)]$.

(iv) *Construction of random matrix* $[\boldsymbol{G}(\beta)]$. Random matrix $[\mathbf{G}(\beta)]$ with values in $\mathbb{M}_{N,n}$ is such that, for all $\ell = 1, \ldots, m$, for all $j_o = 1, \ldots, N_o$, and for all $k = 1, \ldots, n$,

$$[\mathbf{G}(\beta)]_{jk} = [\boldsymbol{\mathcal{G}}(\mathbf{x}^{j_0})]_{\ell k} \quad , \quad j = (\ell, j_o) \in \{1, \ldots, N\} \,, \tag{D.55}$$

in which, for all ℓ in $\{1, \ldots, m\}$ and for all k in $\{1, \ldots, n\}$, the random fields $\{[\boldsymbol{\mathcal{G}}(\mathbf{x})]_{\ell k}, \mathbf{x} \in \mathbb{R}^d\}$ are *independent copies* of a random field $\{\mathcal{Q}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ defined by Eq. (D.38) in which the constant of normalization is chosen such that

$$\zeta = \frac{1}{N} \,. \tag{D.56}$$

From Eqs. (D.27) and (D.55), it can be deduced that the covariance-type matrix $[C_N(\beta)] \in \mathbb{M}_N^+$ introduced in Eq. (3.17) is written, for all ℓ and ℓ' in $\{1, \ldots, m\}$, for all j_o and j'_o in $\{1, \ldots, N_o\}$, and for $j = (\ell, j_o)$ and $j' = (\ell', j'_o)$ in $\{1, \ldots, N\}$, as

$$[C_N(\beta)]_{jj'} = R(\mathbf{x}^{j_o} - \mathbf{x}^{j'_o}; \beta)\, \delta_{\ell\ell'} \,, \tag{D.57}$$

and consequently, $[C_N(\beta)]_{jj} = R(\mathbf{0}, \beta)$ that yields (taking into account Eqs. (D.30), (D.31), and (D.56)):

$$\text{tr}[C_N(\beta)] = 1 \,. \tag{D.58}$$

Let $d[G] = \Pi_{j=1}^N \Pi_{k=1}^n dG_{jk}$ be the volume element on set $\mathbb{M}_{N,n}$ in which dG_{jk} is the Lebesgue measure on $\mathbb{R}$. Taking into account Eqs. (D.38), (D.55), and (D.43), it can easily be proved

that the probability distribution of random matrix $[\mathbf{G}(\beta)]$ admits a probability density function $[G] \mapsto p_{[\mathbf{G}(\beta)]}([G])$ with respect to $d[G]$, which is a mapping from $\mathbb{M}_{N,n}$ into $\mathbb{R}^+$, which is not Gaussian but which is such (using the property of independent copies introduced before and the fact that $\mathcal{Q}_1, \ldots, \mathcal{Q}_d$ are independent Gaussian random processes) that

$$p_{[\mathbf{G}(\beta)]}(-[G]) = p_{[\mathbf{G}(\beta)]}([G]) \quad , \quad \forall\, [G] \in \mathbb{M}_{N,n}\,. \tag{D.59}$$

D.5. Algorithm for the generation of independent realizations of random matrix $[\boldsymbol{G}(\beta)]$

Hereinafter, we detail the algorithm for the generation of independent realizations $[\mathbf{G}(\theta_1;\beta)], \ldots, [\mathbf{G}(\theta_{\nu_s};\beta)]$ of random matrix $[\mathbf{G}(\beta)]$ following the theory presented in Sections D.2 to D.4.

Data:

d: dimension of bounded domain Ω.
m: dimension of the field discretized by the FEM.
N_0: total number of nodes of the FE mesh of domain $\overline{\Omega}$.
$N = m\,N_o$: total number of DOFs before applying the zero Dirichlet conditions.
$\mathbf{x}^1, \ldots, \mathbf{x}^{N_{\text{CD}}}$: N_{CD} nodes of the FE mesh belonging to $\Gamma_0 \subset \partial\Omega$.
$\mathbf{x}^{N_{\text{CD}}+1}, \ldots, \mathbf{x}^{N_o}$: $N_o - N_{\text{CD}}$ nodes of the FE mesh in domain $\Omega \cup \{\partial\Omega \backslash \Gamma_0\}$.
n: reduced-order dimension.
s: real-valued hyperparameter such as $s > 0$.
β: real-valued hyperparameter such as $\beta > 0$.
$[\sigma]$: matrix-valued hyperparameter such that $[\sigma] \in \mathbb{M}_n^{\text{u}}$.
ν_p: sampling points of interval $[-1\,1]$ with $\nu_p = 2\,p$.
ε_r: regularization parameter such that $\varepsilon_r > 0$.
ν_s: number of independent realizations $[\mathbf{G}(\theta_1;\beta)], \ldots, [\mathbf{G}(\theta_{\nu_s};\beta)]$.

Pre-computation:

for $i = 1 : d$
  $L_i = \max_{(j_o, j'_o)} |x_i^{j_o} - x_i^{j'_o}|$.
  $\mathcal{L}_i = \beta\, L_i$.
end

for $\nu = 1 : \nu_p$
  $\kappa_\nu = -1 + \left(\nu - \frac{1}{2}\right)(2/\nu_p)$.
  $\Sigma_\nu = \frac{2}{\nu_p}\,\chi(\kappa_\nu)$ with χ defined by Eq. (D.35).
end

for $j_o = 1 : N_o$
  $\mathbb{1}_\Omega^{\varepsilon_r}(\mathbf{x}^{j_o})$ using Eq. (D.22).
end

Generator:

Initialization of the random generator for uniform random variables.

for $\ell_s = 1 : \nu_s$
  for $k = 1 : n$
    for $\ell = 1 : m$
      Independent realizations of $\Psi_\nu^{i,\ell k}(\theta_{\ell_s})$ on $[0,1]$ and $\Phi_\nu^{i,\ell k}(\theta_{\ell_s})$ on $[0, 2\pi]$ for $i = 1, \ldots, d$

and for $\nu = 1, \ldots, \nu_p$.

for $j_o = 1 : N_o$

$\mathcal{Q}^{1,\ell k}_{j_o}(\theta_{\ell_s}) = \sum_{\nu=1}^{\nu_p} \sqrt{2\,\Sigma_\nu}\,\sqrt{-\log \Psi^{1,\ell k}_\nu(\theta_{\ell_s})}\,\cos\{\Phi^{1,\ell k}_\nu(\theta_{\ell_s}) + \frac{\pi}{\mathcal{L}_1}\kappa_\nu\, x_1^{j_o}\}.$

$\ldots$

$\mathcal{Q}^{d,\ell k}_{j_o}(\theta_{\ell_s}) = \sum_{\nu=1}^{\nu_p} \sqrt{2\,\Sigma_\nu}\,\sqrt{-\log \Psi^{d,\ell k}_\nu(\theta_{\ell_s})}\,\cos\{\Phi^{d,\ell k}_\nu(\theta_{\ell_s}) + \frac{\pi}{\mathcal{L}_d}\kappa_\nu\, x_d^{j_o}\}.$

$\mathcal{G}^{\ell k}_{j_o}(\theta_{\ell_s}) = (1/\sqrt{N}) \times \mathcal{Q}^{\ell k}_{j_o 1}(\theta_{\ell_s}) \times \ldots \times \mathcal{Q}^{\ell k}_{j_o d}(\theta_{\ell_s}).$

$\mathcal{G}^{\varepsilon_r,\ell k}_{j_o}(\theta_{\ell_s}) = \mathbb{1}^{\varepsilon_r}_\Omega(\mathbf{x}^{j_o})\,\mathcal{G}^{\ell k}_{j_o}(\theta_{\ell_s}).$

$[\mathbf{G}(\theta_{\ell_s};\beta)]_{jk} = \mathcal{G}^{\varepsilon_r,\ell k}_{j_o}(\theta_{\ell_s})$ with $j = (\ell, j_o) \in \{1, \ldots, N\}$.

end

end

end

end

Remarks.

- Eq. (D.6) is kept in order to exactly impose the zero Dirichlet condition at the nodes located in Γ_0. Consequently, in the generator given hereinbefore, random field $[\boldsymbol{\mathcal{U}}]$ has been replaced (and it is perfectly licit) by the regularization $[\boldsymbol{\mathcal{G}}^{\varepsilon_r}]$ of $[\boldsymbol{\mathcal{G}}]$.
- The realizations $[\mathbf{G}(\theta_{\ell_s};\beta)]$ for $\ell_s = 1, \ldots, \nu_s$ are not stored in order to avoid core memory problems. The loop in ℓ_s has been introduced for clarifying the sequence related to the initialization of the generator for uniform random variables, but, in practice, this loop is the most external in the code that solves the SROM.

APPENDIX E: DESCRIPTION OF THE LINEAR STATIC SYSTEM USED AS HDM FOR THE EXAMPLE OF SECTION 4

In this appendix, we describe the linear static computational model defined by Eqs. (4.1) and (4.2) for $N = 1\,000$ and $N_{\text{CD}} = 2$. Let $x_1, \ldots, x_N$ be the points in $[0\,,1]$ such that $x_j = (j-1)/(N-1)$. Let $\lambda_1, \ldots, \lambda_N$ be the positive real numbers such that $\lambda_k = 4\pi^2 k^2$, and let $[\lambda]$ be the diagonal matrix in $\mathbb{M}_N^+$ such that $[\lambda]_{kk'} = \lambda_k \delta_{kk'}$. Let $[e]$ be the square matrix in $\mathbb{M}_N$ such that $[e]_{jk} = \sin(k\pi x_j)$. Let $[\Phi] = [\boldsymbol{\varphi}^1 \ldots \boldsymbol{\varphi}^N]$ be the orthogonal matrix in $\mathbb{M}_N$ obtained by the QR decomposition of matrix $[e]$ (we thus have $[e] = [\Phi]\,[R]$ with $[\Phi]\,[\Phi]^T = [\Phi]^T[\Phi] = [I_N]$ and therefore, $< \boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= \delta_{kk'}$) In addition, for $k = 1, \ldots, N$, we have $\varphi^k_1 = [\Phi]_{1k} = 0$ and $\varphi^k_N = [\Phi]_{Nk} = 0$. The computational HDM is then generated as follows:

- Matrix $[K]$ in $\mathbb{M}_N^+$ is written as $[K] = [\Phi]\,[\lambda]\,[\Phi]^T$. Consequently, Eq. (4.1) has a unique solution such that $y_1 = y_N = 0$.
- Vector $\mathbf{f}$ in $\mathbb{R}^N$ is written as $\mathbf{f} = \frac{1}{0.27702}(0.1\,\boldsymbol{\varphi}^2 + 0.4\,\boldsymbol{\varphi}^5 + 0.6\,\boldsymbol{\varphi}^8 + 2.5(\boldsymbol{\varphi}^{29} + \boldsymbol{\varphi}^{30} + \boldsymbol{\varphi}^{31}))$ and consequently, $f_1 = f_N = 0$.
- Let $[b] = [0_{N,N_{\text{CD}}}]$ be the zero matrix in $\mathbb{M}_{N,N_{\text{CD}}}$ except $[b]_{11} = [b]_{N,N_{\text{CD}}} = 1$ with $N_{\text{CD}} = 2$. Matrix $[B]$ in $\mathbb{M}_{N,N_{\text{CD}}}$ such that $[B]^T[B] = [I_{N_{\text{CD}}}]$ is written as the orthonormal basis for the range of $[b]$ (for instance, using Matlab, [B] = orth [b]).

The experimentally measured (surrogate) displacement vector $\mathbf{y}^{\text{exp}}$ that is introduced in Section 4.5 is generated as the solution of $[K]\,\mathbf{y}^{\text{exp}} = \mathbf{f}^{\text{exp}}$ in which the vector $\mathbf{f}^{\text{exp}}$ in $\mathbb{R}^N$ is written as $\mathbf{f}^{\text{exp}} = \frac{1}{0.261466}(0.1\,\boldsymbol{\varphi}^2 + 0.4\,\boldsymbol{\varphi}^5 + 0.6\,\boldsymbol{\varphi}^8 + 2.5(\boldsymbol{\varphi}^{31} + \boldsymbol{\varphi}^{32}) - 0.015\boldsymbol{\varphi}^1))$.

APPENDIX F: ALGORITHM FOR SOLVING THE OPTIMIZATION PROBLEM

The algorithm used for solving the non-convex optimization problem with constraints, defined by Eqs. (4.10) and (5.19), is detailed below. Let $0 < \varepsilon_0$, $0 < \beta_d < \beta_u$, and $\varepsilon_0 < \sigma_u$ (in the applications presented, $\varepsilon_0 = 10^{-2}$, $\beta_d = 0.0005$ to 0.15, $\beta_u = 0.01$ to 0.3, and $\sigma_u = 20$).

- Stage 1. Solving the optimization problem in $\boldsymbol{\alpha} = (s, \beta, [\sigma_{\text{diag}}])$ in which $[\sigma_{\text{diag}}]$ is a diagonal matrix in $\mathbb{M}_n^+$. The number of variables is then $2 + n$. An optimal value $\boldsymbol{\alpha}^{(1)} = (s^{(1)}, \beta^{(1)}, [\sigma_{\text{diag}}^{(1)}])$ of $\boldsymbol{\alpha} = (s, \beta, [\sigma_{\text{diag}}])$ is estimated by solving the optimization problem with the interior-point algorithm with the constraints $\varepsilon_0 \leq s \leq 1$, $\beta_d \leq \beta \leq \beta_u$, and $\varepsilon_0 \leq [\sigma_{\text{diag}}]_{11}, \ldots, [\sigma_{\text{diag}}]_{nn} \leq \sigma_u$. The initial values, which satisfy the constraints, are defined by $\boldsymbol{\alpha}^{(0)} = (s_0, \beta_0, [I_n])$.
- Stage 2. Solving the optimization problem in $[\sigma_{\text{extra}}]$ with $\boldsymbol{\alpha} = (s^{(1)}, \beta^{(1)}, [\sigma_{\text{diag}}^{(1)}] + [\sigma_{\text{extra}}])$ in which $[\sigma_{\text{extra}}]$ is an extra-diagonal upper triangular matrix in $\mathbb{M}_n$. The number of variables is then $n(n-1)/2$. An optimal value $\boldsymbol{\alpha}^{(2)} = (s^{(1)}, \beta^{(1)}, [\sigma^{(2)}])$ of $\boldsymbol{\alpha} = (s, \beta, [\sigma])$, in which $[\sigma^{(2)}] = [\sigma_{\text{diag}}^{(1)}] + [\sigma_{\text{extra}}^{(2)}]$ is estimated using an optimal value $[\sigma_{\text{extra}}^{(2)}]$ of $[\sigma_{\text{extra}}]$, which is estimated by solving the optimization problem using the interior-point algorithm without constraint (only $[\sigma_{\text{extra}}]$ is the variable considered by the optimizer). For $[\sigma_{\text{extra}}]$, the initial condition is $[0_{n,n}]$.
- Stage 3. Let $\beta \mapsto \boldsymbol{\alpha}(\beta) = (s^{(1)}, \beta, [\sigma^{(2)}])$. Stage 3 consists in computing the optimal value $\beta^{\text{opt}} \in [\beta_d\,, \beta_u]$ of β such that $\beta^{\text{opt}} = \min_{\beta \in [\beta_d\,, \beta_u]} J(\boldsymbol{\alpha}(\beta))$ by using the trial method.
- Stage 4. Solving the optimization problem in s and $[\sigma]$ with $\boldsymbol{\alpha} = (s, \beta^{\text{opt}}, [\sigma])$ in which $[\sigma]$ is an upper triangular matrix in $\mathbb{M}_n^u$. The number of variables is then $1 + n(n+1)/2$. An optimal value $\boldsymbol{\alpha}^{\text{opt}} = (s^{\text{opt}}, \beta^{\text{opt}}, [\sigma^{\text{opt}}])$ of $\boldsymbol{\alpha} = (s, \beta^{\text{opt}}, [\sigma])$ is estimated by solving the optimization problem with the interior-point algorithm with the constraints $\varepsilon_0 \leq s \leq 1$ and $\varepsilon_0 \leq [\sigma]_{11}, \ldots, [\sigma]_{nn} \leq \sigma_u$ (only s and $[\sigma]$ are the variables considered by the optimizer). The initial condition is $\boldsymbol{\alpha}^{(0)} = (s^{(1)}, \beta^{\text{opt}}, [\sigma^{(2)}])$.

APPENDIX G: CONSTRUCTION OF THE DAMPING MATRIX $[\mathbb{D}]$

The construction proposed for $[\mathbb{D}]$ corresponds to an adaptation of the representation presented in [71] (Eq. (45) of Chapter VI). Let $\{\boldsymbol{\varphi}^1, \ldots, \boldsymbol{\varphi}^n\}$ be the first n elastic modes associated with the first n eigenvalues $0 < \lambda_1 < \ldots < \lambda_n$ of the linear undamped structure associated with the nonlinear damped dynamical system and computed in solving Eqs. (5.10) and (5.11). Matrix $[\mathbb{D}]$ belonging to $\mathbb{M}_N^+$ is written as

$$[\mathbb{D}] = \sum_{k=1}^{n} 2\,\xi_d\,(\sqrt{\lambda_k} - \sqrt{\lambda_n})\,[M]\,\boldsymbol{\varphi}^k\,([M]\,\boldsymbol{\varphi}^k)^T + 2\,\xi_d\,\sqrt{\lambda_n}\,[M]\,. \tag{G.1}$$

Using Eq. (5.12), it can be seen that

$$[V]^T[\mathbb{D}]\,[V] = [\mathbb{D}^{(n)}]\,, \tag{G.2}$$

in which the reduced damping matrix $[\mathbb{D}^{(n)}] \in \mathbb{M}_n^+$ is diagonal and for which the diagonal entries are $[\mathbb{D}^{(n)}]_{kk} = 2\,\xi_d\,\sqrt{\lambda_k}$. For $\{k > n\,, k' > n\}$, we have $< [\mathbb{D}]\boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= 2\,\xi_d\,\sqrt{\lambda_n}\delta_{kk'}$. For $\{k \leq n\,, k' > n\}$ or for $\{k > n\,, k' \leq n\}$, we have $< [\mathbb{D}]\boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= 0$. The model defined by Eq. (G.1) allows for obtaining the constant damping rate ξ_d for the first n elastic modes and for obtaining the damping rates $\{\xi_d\,\sqrt{\lambda_n}/\sqrt{\lambda_{n+1}}, \ldots, \xi_d\,\sqrt{\lambda_n}/\sqrt{\lambda_N}\}$ for the $N - n$ elastic modes $\{\boldsymbol{\varphi}^{n+1}, \ldots, \boldsymbol{\varphi}^N\}$ (that are not computed).

The assemblage of the full matrix $[\mathbb{D}]$ defined by Eq. (G.1) is never done.

- Let $\mathbf{z}$ be a vector in $\mathbb{R}^N$. We then have

$$[\mathbb{D}]\,\mathbf{z} = \sum_{k=1}^{n} 2\,\xi_d\,(\sqrt{\lambda_k} - \sqrt{\lambda_n})\,[M]\,\boldsymbol{\varphi}^k\; < [M]\,\boldsymbol{\varphi}^k\,, \mathbf{z} > + 2\,\xi_d\,\sqrt{\lambda_n}\,[M]\,\mathbf{z}\,. \tag{G.3}$$

This equation shows that the complexity of the product of matrix $[\mathbb{D}]$ by vector $\mathbf{z}$ is induced by the products $[M]\,\boldsymbol{\varphi}^1, \ldots, [M]\,\boldsymbol{\varphi}^n, [M]\,\mathbf{z}$, that is to say by $n+1$ products of the sparse matrix $[M]$ by a vector.

- Let $[\mathbf{W}]$ be the SROB with values in $\mathbb{M}_{N,n}$ and verifying Eq. (2.12) and Eq. (2.11). We then have

$$[\mathbf{W}]^T[\mathbb{D}]\,[\mathbf{W}] = \sum_{k=1}^{n} 2\,\xi_d\,(\sqrt{\lambda_k} - \sqrt{\lambda_n})\,\boldsymbol{\Psi}^k(\boldsymbol{\Psi}^k)^T + 2\,\xi_d\,\sqrt{\lambda_n}\,[I_n]\,, \tag{G.4}$$

in which $\boldsymbol{\Psi}^k = [\mathbf{W}]^T[M]\,\boldsymbol{\varphi}^k$ is a random vector with values in $\mathbb{R}^n$.

APPENDIX H: CONSTRUCTION OF THE COST FUNCTION FOR THE IDENTIFICATION OF THE HYPERPARAMETER IN THE APPLICATION PRESENTED IN SECTION 4

For constructing the cost function, we introduce the random function $\omega \mapsto \mathbf{dB}^{(n)}(\omega;\boldsymbol{\alpha}) = (\mathrm{dB}_1^{(n)}(\omega;\boldsymbol{\alpha}), \ldots, \mathrm{dB}_{m_o}^{(n)}(\omega;\boldsymbol{\alpha}))$ defined on $\mathcal{B}_o$ with valued in $\mathbb{R}^{m_o}$ such that, for all $j = 1, \ldots, m_o$ (with $m_o = 122$),

$$\mathrm{dB}_j^{(n)}(\omega;\boldsymbol{\alpha}) = \log_{10}(|\widehat{O}_j^{(n)}(\omega;\boldsymbol{\alpha})|)\,, \tag{H.1}$$

in which $\widehat{\mathbf{O}}^{(n)}(\omega;\boldsymbol{\alpha}) = (\widehat{O}_1^{(n)}(\omega;\boldsymbol{\alpha}), \ldots, \widehat{O}_{m_o}^{(n)}(\omega;\boldsymbol{\alpha}))$ is defined by Eq. (5.18).

In order to define the target functions for constructing the cost function, we introduce the functions $\omega \mapsto \mathbf{db}^{\mathrm{ref}}(\omega) = (\mathrm{db}_1^{\mathrm{ref}}(\omega), \ldots, \mathrm{db}_{m_o}^{\mathrm{ref}}(\omega))$ and $\omega \mapsto \mathbf{db}^{(n)}(\omega) = (\mathrm{db}_1^{(n)}(\omega), \ldots, \mathrm{db}_{m_o}^{(n)}(\omega))$, defined on $\mathcal{B}_o$ with values in $\mathbb{R}^{m_o}$ such that, for all $j = 1, \ldots, m_o$,

$$\mathrm{db}_j^{\mathrm{ref}}(\omega) = \log_{10}(|\widehat{o}_j(\omega)|) \quad , \quad \mathrm{db}_j^{(n)}(\omega) = \log_{10}(|\widehat{o}_j^{(n)}(\omega)|)\,, \tag{H.2}$$

in which $\widehat{\mathbf{o}}(\omega) = (\widehat{o}_1(\omega), \ldots, \widehat{o}_{m_o}(\omega))$ is defined by Eq. (5.8) and where $\widehat{\mathbf{o}}^{(n)}(\omega) = (\widehat{o}_1^{(n)}(\omega), \ldots, \widehat{o}_{m_o}^{(n)}(\omega))$ is defined by Eq. (5.15). The cost function, defined by Eq. (2.17), is rewritten here as

$$J(\boldsymbol{\alpha}) = w_J\,J_{\mathrm{mean}}\,(\boldsymbol{\alpha}) + (1 - w_J)\,J_{\mathrm{std}}\,(\boldsymbol{\alpha})\,, \tag{H.3}$$

in which

$$J_{\mathrm{mean}}\,(\boldsymbol{\alpha}) = \frac{1}{c_{\mathrm{mean}}} \sum_{j=1}^{m_o} \int_{\mathcal{B}_o} |\mathrm{db}_j^{\mathrm{ref}}(\omega) - E\{\mathrm{dB}_j^{(n)}(\omega;\boldsymbol{\alpha})\}|^2\,w_j(\omega)\,d\omega\,, \tag{H.4}$$

$$J_{\mathrm{std}}\,(\boldsymbol{\alpha}) = \frac{1}{c_{\mathrm{std}}} \sum_{j=1}^{m_o} \int_{\mathcal{B}_o} |v_j^{(\mathrm{ref},n)}(\omega) - v_j^{(n)}(\omega;\boldsymbol{\alpha})|^2\,w_j(\omega)\,d\omega\,, \tag{H.5}$$

in which the positive constants c_{mean} and c_{std} are defined by

$$c_{\mathrm{mean}} = \sum_{j=1}^{m_o} \int_{\mathcal{B}_o} |\mathrm{db}_j^{\mathrm{ref}}(\omega)\}|^2\,w_j(\omega)\,d\omega \quad , \quad c_{\mathrm{std}} = \sum_{j=1}^{m_o} \int_{\mathcal{B}_o} |v_j^{(\mathrm{ref},n)}(\omega)|^2\,w_j(\omega)\,d\omega\,. \tag{H.6}$$

In these equations, $\mathbf{v}^{(\mathrm{ref},n)}(\omega) = (v_1^{(\mathrm{ref},n)}(\omega), \ldots, v_{m_o}^{(\mathrm{ref},n)}(\omega))$ and $\mathbf{v}^{(n)}(\omega;\boldsymbol{\alpha}) = (v_1^{(n)}(\omega;\boldsymbol{\alpha}), \ldots, v_{m_o}^{(n)}(\omega;\boldsymbol{\alpha}))$ are defined, for $j = 1, \ldots m_o$, by

$$v_j^{(\mathrm{ref},n)}(\omega) = \gamma\,|\mathrm{db}_j^{\mathrm{ref}}(\omega) - \mathrm{db}_j^{(n)}(\omega)|\,, \tag{H.7}$$

in which $\gamma > 0$ is an amplitude factor and where

$$v_j^{(n)}(\omega;\boldsymbol{\alpha}) = \{\,E\{dB_j^{(n)}(\omega;\boldsymbol{\alpha})^2\} - (E\{dB_j^{(n)}(\omega;\boldsymbol{\alpha})\})^2\}\,\}^{1/2}\,. \tag{H.8}$$

In Eqs. (H.4) to (H.6), for $j = 1, \ldots m_o$, the function $\omega \mapsto w_j(\omega)$ are bounded on $\mathcal{B}_o$ with values in $\mathbb{R}^+$. In the application presented in Section 4, these functions are chosen such that $w_j(\omega) = |\mathrm{db}_j^{\mathrm{ref}}(\omega) - \mathrm{db}_j^{(n)}(\omega)|$ for all ω in $\mathcal{B}_o$.

APPENDIX I: DESCRIPTION OF THE μ-PARAMETRIC HDM FOR THE NUMERICAL EXAMPLE OF SECTION 6

In this appendix, we describe the μ-parametric HDM, which is the linear static computational model defined by Eqs. (6.1) and (6.2) for $N = 1\,000$ and $N_{\text{CD}} = 2$. The common part of the model described in Appendix E is rewritten below in order to facilitate the readability of the paper.

Let $x_1, \ldots, x_N$ be the points in $[0\,,1]$ such that $x_j = (j-1)/(N-1)$. Let $\lambda_1, \ldots, \lambda_N$ be the positive real numbers such that $\lambda_k = 4\pi^2 k^2$, and let $[\lambda]$ be the diagonal matrix in $\mathbb{M}_N^+$ such that $[\lambda]_{kk'} = \lambda_k \delta_{kk'}$. Let $[e]$ be the square matrix in $\mathbb{M}_N$ such that $[e]_{jk} = \sin(k\pi x_j)$. Let $[\Phi] = [\boldsymbol{\varphi}^1 \ldots \boldsymbol{\varphi}^N]$ be the orthogonal matrix in $\mathbb{M}_N$ obtained by the QR decomposition of matrix $[e]$ (we thus have $[e] = [\Phi]\,[R]$ with $[\Phi]\,[\Phi]^T = [\Phi]^T[\Phi] = [I_N]$ and therefore, $< \boldsymbol{\varphi}^k\,, \boldsymbol{\varphi}^{k'} >= \delta_{kk'}$) In addition, for $k = 1, \ldots, N$, we have $\varphi_1^k = [\Phi]_{1k} = 0$ and $\varphi_N^k = [\Phi]_{Nk} = 0$. The computational HDM is then generated as follows:

- Matrix $[K]$ in $\mathbb{M}_N^+$ is written as $[K] = [\Phi]\,[\lambda]\,[\Phi]^T$. Consequently, Eq. (6.1) has a unique solution such that $y_1 = y_N = 0$.
- For μ in $\mathcal{C}_{\boldsymbol{\mu}} = [\,\mu_{\min}\,, \mu_{\max}\,] \subset \mathbb{R}$, the vector $\mathbf{f}(\mu) = (f_1(\mu), \ldots, f_N(\mu))$ in $\mathbb{R}^N$ is written as $\mathbf{f}(\mu) = g_{\max}(\mu)^{-1}\,\mathbf{g}(\mu)$ in which the vector $\mathbf{g}(\mu) = (g_1(\mu), \ldots, g_N(\mu))$ in $\mathbb{R}^N$ is written as $\mathbf{g}(\mu) = 0.1\,\boldsymbol{\varphi}^2 + \mu\,\boldsymbol{\varphi}^3 + 0.4\,\boldsymbol{\varphi}^5 + 0.6\,\boldsymbol{\varphi}^8$ and where $g_{\max}(\mu) = \max_{j=1,\ldots,N} g_j(\mu)$. Consequently, we have $f_1(\mu) = f_N(\mu) = 0$.
- Let $[b] = [0_{N,N_{\text{CD}}}]$ be the zero matrix in $\mathbb{M}_{N,N_{\text{CD}}}$ except $[b]_{11} = [b]_{N,N_{\text{CD}}} = 1$ with $N_{\text{CD}} = 2$. Matrix $[B]$ in $\mathbb{M}_{N,N_{\text{CD}}}$ such that $[B]^T[B] = [I_{N_{\text{CD}}}]$ is written as the orthonormal basis for the range of $[b]$ (for instance, using Matlab, [B] = orth [b]).

REFERENCES

1. International Assessment of Research and Development in Simulation-based Engineering and Science, *http://www.wtec.org/sbes/*
2. Grepl MA, Maday Y, Nguyen NC, Patera A. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. ESAIM: Mathematical Modelling and Numerical Analysis 2007; 41(03):575-605.
3. Nguyen N, Peraire J. An efficient reduced-order modeling approach for nonlinear parametrized partial differential equations. International Journal for Numerical Methods in Engineering 2008; 76(1):27-55.
4. Chaturantabut S, Sorensen DC. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific and Statistical Computing* 2010; **32**(5):2737-2764.
5. Degroote J, Virendeels J, Willcox K. Interpolation among reduced-order matrices to obtain parameterized models for design, optimization and probabilistic analysis, *International Journal for Numerical Methods in Fluids* 2010; **63**:207-230.
6. Carlberg K, Bou-Mosleh C, Farhat C. Efficient nonlinear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations. *International Journal for Numerical Methods in Engineering* 2011; **86**(2):155-181.
7. Carlberg K, Farhat C. A low-cost, goal-oriented compact proper orthogonal decomposition basis for model reduction of static systems. *International Journal for Numerical Methods in Engineering* 2011; **86**(3):381-402.
8. Amsallem D, Zahr MJ, Farhat C. Nonlinear model order reduction based on local reduced-order bases. *International Journal for Numerical Methods in Engineering* 2012; **92**(10):891-916.
9. Carlberg K, Farhat C, Cortial J, Amsallem D. The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows. *Journal of Computational Physics* 2013; **242**:623-647.
10. Zahr M, Farhat C. Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *International Journal for Numerical Methods in Engineering* 2015; *102*(5):1077-1110.
11. Amsallem D, Zahr M, Choi Y, Farhat C. Design optimization using hyper-reduced-order models. *Structural and Multidisciplinary Optimization* 2015; 51(4):919-940.
12. Paul-Dubois-Taine A, Amsallem D. An adaptive and efficient greedy procedure for the optimal training of parametric reduced-order models. *International Journal for Numerical Methods in Engineering* 2015; **102**(5):1262-1292.
13. Farhat C, Avery P, Chapman T, Cortial J. Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. *International Journal for Numerical Methods in Engineering* 2014; *98*(9):625-662.
14. Farhat C, Chapman T, Avery P. Structure-preserving, stability, and accuracy properties of the Energy-Conserving Sampling and Weighting (ECSW) method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering* 2015; 102(5):1077-1110.
15. Ryckelynck D. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics* 2005; **202**:346-366.

16. Beck JL, Katafygiotis LS. Updating models and their uncertainties - I: Bayesian statistical framework. *Journal of Engineering Mechanics* 1998; **124**(4):455-461.
17. Ghanem R, Spanos PD. *Stochastic Finite Elements: a Spectral Approach*. Springer-Verlag: New York, 1991.
18. Ghanem R, Spanos PD. *Stochastic Finite Elements: A spectral Approach (revised edition)*. Dover Publications: New York, 2003.
19. Soize C, Ghanem R. Physical systems with random uncertainties: Chaos representation with arbitrary probability measure. *SIAM Journal On Scientific Computing* 2004; **26**(2): 395-410.
20. Mace R, Worden W, Manson G. Uncertainty in Structural Dynamics.*Special issue of the Journal of Sound and Vibration* 2005; **288**(3):431-790.
21. Schueller GI. Computational Methods in Stochastic Mechanics and Reliability Analysis. *Special issue of Computer Methods in Applied Mechanics and Engineering* 2005; **194**(12-16):1251-1795.
22. Schueller GI. Uncertainties in Structural Mechanics and Analysis-Computational Methods. *Special issue of Computer and Structures* 2005; **83**(14):1031-1150.
23. Schueller GI. Developments in stochastic structural mechanics. *Archive of Applied Mechanics* 2006; **75**(10-12):755-773.
24. Deodatis G, Spanos PD. 5th International Conference on Computational Stochastic Mechanics. *Special issue of the Probabilistic Engineering Mechanics* 2008; **23**(2-3):103-346.
25. Schueller GI, Pradlwarter HJ. Uncertain linear systems in dynamics: Retrospective and recent developments by stochastic approaches. *Engineering Structures* 2009; **31**(11):2507-2517.
26. Le Maitre OP, Knio OM. *Spectral Methods for Uncerainty Quantification with Applications to Computational Fluid Dynamics*. Springer: Heidelberg, 2010.
27. Soize C. *Stochastic Models of Uncertainties in Computational Mechanics*. American Society of Civil Engineers (ASCE): Reston, 2012.
28. Bui-Thanh T, Willcox K, Ghattas O. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamic applications. *AIAA Journal* 2008; **46**(10):2520-2529.
29. Soize C. Stochastic modeling of uncertainties in computational structural dynamics - Recent theoretical advances. *Journal of Sound and Vibration* 2013; **332**(10): 2379-2395.
30. Marzouk YM, Najm HN, Rahn LA. Stochastic spectral methods for efficient Bayesian solution of inverse problems.*Journal of Computational Physics* 2007; **224**(2):560-586.
31. Galbally D, Fidkowski K, Willcox K, Ghattas O. Non-linear model reduction for uncertainty quantification in large scale inverse problems. *International Journal for Numerical Methods in Engineering* 2010; **81**(12):1581-1608.
32. Lieberman C, Willcox K, Ghattas O. Parameter and state model reduction for large scale statistical inverse problems, *SIAM Journal on Scientific Computing* 2010; **32**:2523-2542.
33. Soize C. Identification of high-dimension polynomial chaos expansions with random coefficients for non-Gaussian tensor-valued random fields using partial and limited experimental data. *Computer Methods in Applied Mechanics and Engineering* 2010; **199**(33-36):2150-2164.
34. Soize C. A computational inverse method for identification of non-Gaussian random fields using the Bayesian approach in very high dimension. *Computer Methods in Applied Mechanics and Engineering* 2011; **200**(45-46):3083-3099.
35. Nouy A, Soize C. Random fields representations for stochastic elliptic boundary value problems and statistical inverse problems. *European Journal of Applied Mathematics* 2014; **25**(3):339-373.
36. Cui T, Marzouk YM, Willcox KE. Data-driven model reduction for the Bayesian solution of inverse problems. *International Journal for Numerical Methods in Engineering* 2015; **102**(5):966-990.
37. Soize C. Random vectors and random fields in high dimension. Parametric model-based representation, identification from data, and inverse problems. pp:1-65, in *Handbook for Uncertainty Quantification*, edited by Ghanem R, Higdon D, Owhadi H., Springer: Heidelberg, 2016.
38. Soize C. A nonparametric model of random uncertainties for reduced matrix models in structural dynamics. *Probabilistic Engineering Mechanics* 2000; **15**(3):277-294.
39. Soize C. Maximum entropy approach for modeling random uncertainties in transient elastodynamics, *Journal of the Acoustical Society of America* 2001; **109**(5):1979-1996.
40. Soize C. Random matrix theory for modeling random uncertainties in computational mechanics. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**(12-16):1333-1366.
41. Soize C. Random Matrix Models and Nonparametric Method for Uncertainty Quantification. pp:1-84, *Handbook for Uncertainty Quantification*, edited by Ghanem R, Higdon D, Owhadi H., Springer: Heidelberg, 2016.
42. Shannon CE. A mathematical theory of communication. *Bell System Technical Journal* 1948; **27**:379-423 and 623-659.
43. Jaynes ET. Information theory and statistical mechanics. *Physical Review* 1957; **106**(4):620-630 and **108**(2):171-190.
44. Kapur JN, Kesavan HK.*Entropy Optimization Principles with Applications*. Academic Press: San Diego, 1992.
45. Mignolet MP, Soize C. Nonparametric stochastic modeling of linear systems with prescribed variance of several natural frequencies. *Probabilistic Engineering Mechanics* 2008; **23**(2-3):267-278.
46. Soize C. Nonparametric probabilistic approach of uncertainties for elliptic boundary value problem.*International Journal for Numerical Methods in Engineering* 2009; **80**(6-7):673-688.
47. Chen C, Duhamel D, Soize C. Probabilistic approach for model and data uncertainties and its experimental identification in structural dynamics: Case of composite sandwich panels. *Journal of Sound and Vibration* 2006; **294**(1-2):64-81.
48. Ohayon R, Soize C. *Advanced Computational Vibroacoustics. Reduced-Order Models and Uncertainty Quantification*. Cambridge University Press: Cambridge, 2014.
49. Soize C, Poloskov IE. Time-domain formulation in computational dynamics for linear viscoelastic media with model uncertainties and stochastic excitation. *Computers and Mathematics with Applications* 2012; **64**(11):3594-3612.

50. Soize C, Chebli H. Random uncertainties model in dynamic substructuring using a nonparametric probabilistic model. *Journal of Engineering Mechanics* 2003; **129**(4):449-457.
51. Chebli H, Soize C. Experimental validation of a nonparametric probabilistic model of non homogeneous uncertainties for dynamical systems. *Journal of the Acoustical Society of America* 2004; **115**(2):697-705.
52. Capiez-Lernout E, Soize C, Lombard JP, Dupont C, Seinturier E. Blade manufacturing tolerances definition for a mistuned industrial bladed disk. *Journal of Engineering for Gas Turbines and Power* 2005; **127**(3):621-628.
53. Duchereau J, Soize C. Transient dynamics in structures with nonhomogeneous uncertainties induced by complex joints. *Mechanical Systems and Signal Processing* 2006; **20**(4):854-867.
54. Mignolet MP, Soize C, Avalos J. Nonparametric stochastic modeling of structures with uncertain boundary conditions / coupling between substructures. *AIAA Journal* 2013; **51**(6):1296-1308.
55. Durand JF, Soize C, Gagliardini L. Structural-acoustic modeling of automotive vehicles in presence of uncertainties and experimental identification and validation. *Journal of the Acoustical Society of America* 2008; **124**(3):1513-1525.
56. Fernandez C, Soize C, Gagliardini L. Sound-insulation layer modelling in car computational vibroacoustics in the medium-frequency range. *Acta Acustica United with Acustica (AAUWA)* 2010; **96**(3):437-444.
57. Ohayon R, Soize C. Advanced computational dissipative structural acoustics and fluid-structure interaction in low- and medium-frequency domains. Reduced-order models and uncertainty quantification. *International Journal of Aeronautical and Space Sciences* 2012; **13**(2):127-153.
58. Desceliers C, Soize C, Cambier S. Non-parametric-parametric model for random uncertainties in nonlinear structural dynamics - Application to earthquake engineering. *Earthquake Engineering and Structural Dynamics* 2004; **33**(3):315-327.
59. Cottereau R, Clouteau D, Soize C. Construction of a probabilistic model for impedance matrices. *Computer Methods in Applied Mechanics and Engineering* 2007; **196**(17-20):2252-2268.
60. Cottereau R, Clouteau D, Soize C. Probabilistic impedance of foundation: impact of the seismic design on uncertain soils. *Earthquake Engineering and Structural Dynamics* 2008; **37**(6):899-918.
61. Capiez-Lernout E, Soize C. Design optimization with an uncertain vibroacoustic model. *Journal of Vibration and Acoustics* 2008; **130**(2):1-8.
62. Capiez-Lernout E, Soize C. Robust design optimization in computational mechanics. *Journal of Applied Mechanics - Transactions of the ASME* 2008; **75**(2):1-11.
63. Mignolet MP, Soize C. Stochastic reduced order models for uncertain nonlinear dynamical systems. *Computer Methods in Applied Mechanics and Engineering* 2008; **197**(45-48):3951-3963.
64. Capiez-Lernout E, Soize C, Mignolet MP. Post-buckling nonlinear static and dynamical analyses of uncertain cylindrical shells and experimental validation. *Computer Methods in Applied Mechanics and Engineering* 2014; **271**(1):210-230.
65. Edelman A, Arias TA, Smith ST. The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications* 1998; **20**(2):303-353.
66. Absil PA, Mahony R, Sepulchre R. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press: Princeton, 2008.
67. Golub GH, Van Loan CF, *Matrix Computations (Third Edition)*. The Johns Hopkins University Press: Baltimore, 1996.
68. Shinozuka M, Wen YK. Monte Carlo solution of nonlinear vibration. *AIAA Journal* 1972; **10**(1):37-40.
69. Shinozuka M, Deodatis G. Response variability of stochastic finite element systems. *Journal of Engineering Mechanics*, 1988; **114**(3):499-519.
70. Poirion F, Soize C. Numerical methods and mathematical aspects for simulation of homogeneous and non homogeneous Gaussian vector fields, pp. 17-53 in *Probabilistic Methods in Applied Physics*, P. Krée and W. Wedig (Eds.), *Lecture Notes in Physics*, ISBN 3-540-60214-3, Springer-Verlag.
71. Ohayon R, Soize C. *Structural Acoustics and Vibration*. Academic Press: San Diego, London, 1998.